DOCUMENT RESUME

ED 214 226 CS 503 794

TITLE	Speech Research: A Report on the Status and Progress of Studies on the Nature of Speech, Instrumentation for Its Investigation, and Practical Applications, January 1-March 31, 1982.
INSTITUTION	Haskins Labs., New Haven, Conn.
SPONS AGENCY	National Inst. of Child Health and Human Development (NIH), Bethesda, Md.; National Inst. of Neurological and Communicative Disorders and Stroke (NIH), Bethesda, Md.; National Science Foundation, Washington, D.C.
REPORT NO	SR-69
PUB DATE	82
CONTRACT	NICHHD-N01-HD-1-2420
GRANT	NICHHD-HD-01994; NIH-RR-05596; NINCDS-NS-13870; NSF-PRF-8006144
NOTE	301p.
EDRS PRICE	MF01/PC13 Plus Postage.
DESCRIPTORS	*Articulation (Speech); Auditory Perception; *Hearing Impairments; Language Skills; Memory; *Oral Language; Perceptual Motor Coordination; Psychomotor Skills; Reading Ability; Sign Language; *Speech Communication; *Speech Handicaps; Speech Skills

ABSTRACT

One of a regular series on the status and progress of studies into the nature of speech, instrumentation for its investigation, and practical applications of research, this report covers the period of January 1 to March 31, 1982. The 13 studies deal with the following topics: (1) speech perception and memory coding in relation to reading ability, (2) the use of orthographic structure by deaf adults, (3) information support for speech, (4) the stream of speech, (5) using the acoustic signal to make inferences about place and duration of tongue-palate contact, (6) the patterns of human interlimb coordination that have emerged from the properties of nonlinear limit cycle oscillatory processes, (7) motor control, (8) the nature of motor control in Down's syndrome, (9) periodicity and auditory memory, (10) reading skill and language skill, (11) the role of sign order and morphological structure in memory for American Sign Language sentences, (12) perception of nasal consonants with special reference to Catalan, and (13) speech production characteristics of the hearing impaired. (FL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED214226

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

SR-69 (1982)

Status Report on

SPEECH RESEARCH

A Report on
the Status and Progress of Studies on
the Nature of Speech, Instrumentation
for its Investigation, and Practical
Applications

1 January - 31 March 1982

Haskins Laboratories
270 Crown Street
New Haven, Conn. 06510

Distribution of this document is unlimited.

(This document contains no information not freely available to the general public. Haskins Laboratories distributes it primarily for library use. Copies are available from the National Technical Information Service or the ERIC Document Reproduction Service. See the Appendix for order numbers of previous Status Reports.)

CS 503 794

ACKNOWLEDGMENTS

The research reported here was made possible in part by support from the following sources:

National Institute of Child Health and Human Development
Grant HD-01994
Grant HD-16591

National Institute of Child Health and Human Development
Contract N01-HD-1-2420

National Institutes of Health
Biomedical Research Support Grant RR-05596

National Science Foundation
Grant PRF-8006144
Grant BNS-8111470

National Institute of Neurological and Communicative Disorders and Stroke
Grant NS13870
Grant NS13617
Grant NS18010

HASKINS LABORATORIES

Personnel in Speech Research

Alvin M. Liberman,* President and Research Director
Franklin S. Cooper,* Associate Research Director
Patrick W. Nye, Associate Research Director
Raymond C. Huey,* Treasurer
Alice Dadourian, Secretary

Investigators

Arthur S. Abramson*
Peter J. Alfonso*
Thomas Baer
Alan Bell[1]
Fredericka Bell-Berti*
Catherine Best*
Gloria J. Borden*
Susan Brady*
Giuseppe Cossu[2]
Robert Crowder*
Carol A. Fowler*
Louis Goldstein*
Vicki L. Hanson
Katherine S. Harris*
Alice Healy*
Kiyoshi Honda[3]
Leonard Katz*
J. A. Scott Kelso
Andrea G. Levitt*
Isabelle Y. Liberman*
Leigh Lisker*
Virginia Mann*
Charles Marshall
Ignatius G. Mattingly*
Nancy S. McGarr*
Lawrence J. Raphael*
Bruno H. Repp
Philip E. Rubin
Elliot Saltzman
Donald P. Shankweiler*
Michael Studdert-Kennedy*
Betty Tuller*
Michael T. Turvey*
Robert Verbrugge*

Technical and Support Staff

Eric L. Andreasson
Margo Carter
Elizabeth P. Clark
Vincent Gulisano
Donald Hailey
Sabina D. Koroluk
Bruce Martin
Agnes M. McKeon*
Nancy O'Brien
Marilyn K. Parnell
William P. Scully
Richard S. Sharkany
Leonard Szubowicz
Edward R. Wiley
Mary-Anne Wolf
David Zeichner

Students*

Suzanne Boyce
Andre Cooper
Tova Clayman
Steven Eady
Jo Estill
Laurie B. Feldman
Nancy Fishbein
Carole E. Gelfer
Janette Henderson
Charles Hoequist
Robert Katz
Rena Krakow
Peter Kugler
Gerald Lame
Anthony Levas
Harriet Magen
Sharon Manuel
Richard McGowan
Suzi Pollock
Brad Rakerd
Daniel Recasens
Rosemarie Rotunno
Hyla Rubin
Judith Rubin
Arnold Shapiro
Suzanne Smith
Ben C. Watson
Douglas Whalen
Deborah Wilkenfeld
David Williams

---

*Part-time
[1] Visiting from University of Colorado, Boulder, Colorado
[2] Visiting from Istituto Di Neuropsichiatria Infantile, Sassari, Italy
[3] Visiting from University of Tokyo, Japan

SR-69 (1982)
(January-March)

CONTENTS

I. Manuscripts and Extended Reports

Speech perception and memory coding in relation to reading ability--Susan Brady, Donald Shankweiler, and Virginia Mann 1

The use of orthographic structure by deaf adults: Recognition of fingerspelled letters--Vicki L. Hanson 27

Exploring the information support for speech--J. A. Scott Kelso and Betty Tuller 43

The stream of speech--Robert E. Remez and Philip E. Rubin 55

Using the acoustic signal to make inferences about place and duration of tongue-palate contact--P. J. Price 61

Patterns of human interlimb coordination emerge from the properties of non-linear limit cycle oscillatory processes: Theory and data--J. A. Scott Kelso, Kenneth G. Holt, Philip Rubin, and Peter N. Kugler 71

Motor control: Which themes do we orchestrate?--J. A. S. Kelso and E. L. Saltzman 111

Exploring the nature of motor control in Down's syndrome--Walter E. Davis and J. A. Scott Kelso 123

Periodicity and auditory memory: A pilot study--Janet May and Bruno H. Repp 145

Reading skill and language skill--Virginia A. Mann 151

On the role of sign order and morphological structure in memory for American Sign Language sentences--Vicki L. Hanson and Ursula Bellugi 171

Perception of nasal consonants with special reference to Catalan--Daniel Recasens 189

Speech production characteristics of the hearing impaired--Mary Joe Osberger and Nancy S. McGarr 227

II. Publications 291

III. Appendix: DTIC and ERIC numbers (SR-21/22 - SR-66) 293

I. MANUSCRIPTS AND EXTENDED REPORTS

SPEECH PERCEPTION AND MEMORY CODING IN RELATION TO READING ABILITY

Susan Brady, Donald Shankweiler,+ and Virginia Mann++

Abstract. Previous work has demonstrated that children who are poor readers have short-term memory deficits in tasks in which the stimuli lend themselves to phonetic coding. The aim of the present study was to explore whether the poor readers' memory deficit may have its origin in perception with the encoding of the stimuli. Three experiments were conducted with third-grade good and poor readers. As in earlier experiments, the poor readers were found to perform less well on recall of random word strings and to be less affected by the phonetic characteristics (rhyming or not rhyming) of the items (Experiment 1). In addition, the poor readers produced more errors of transposition (in the nonrhyming strings) than did the good readers, a further indication of the poor readers' problems with memory for order. The subjects were tested on two auditory perception tasks, one employing words (Experiment 2) and the other nonspeech environmental sounds (Experiment 3). Each was presented under two conditions: with a favorable signal-to-noise ratio and with masking. The poor readers made significantly more errors than the good readers when listening to speech in noise, but did not differ in perception of speech without noise or in perception of nonspeech environmental sounds, whether noise-masked or not. Together, the results of the perception studies suggest that poor readers have a perceptual difficulty that is specific to speech. It is suggested that the short-term memory deficits characteristic of poor readers may stem from material-specific problems of perceptual processing.

+Also University of Connecticut, Storrs, Ct.
++Also Bryn Mawr College, Bryn Mawr, Pa.

Acknowledgment. The authors wish to thank several colleagues for their helpful comments and suggestions: Carol Fowler, Vicki Hanson, Leonard Katz, Robert Katz, and Alvin Liberman. We are especially grateful to Isabelle Liberman, whose insightful comments on earlier drafts contributed to the clarity of this presentation. We are also indebted to William Holland, Superintendent of Schools, Narragansett, Rhode Island, and to the faculty at the Narragansett Elementary School for their kind cooperation: David Hayes, Principal; Judy Aiello, Reading Coordinator; and the third grade teachers (Sue Boland, Leslie Flynn, Edward O'Brien, Hope Rawlings, and Marguerite Strain). This research was supported by a grant from the National Institute of Child Health and Human Development (Grant HD-01994) to Haskins Laboratories.

[HASKINS LABORATORIES: Status Report on Speech Research SR-69 (1982)]

Many studies have shown that children who are poor readers tend to perform deficiently on short-term memory tasks. There is considerable evidence, however, that the memory problem is specific to linguistic material and to other material that lends itself to linguistic representation. A hypothesis has been proposed that failure to make effective use of phonetic coding in short-term memory may account for some of the deficiencies poor readers typically show in language processing (Liberman, Shankweiler, Liberman, Fowler, & Fischer, 1977). Tests of this hypothesis have utilized the well-known phenomenon that when normal adult subjects are required to recall strings of rhyming and nonrhyming letters or words, many more errors typically occur on the rhyming strings (Baddeley, 1966; Conrad, 1964, 1972). Children who are good readers, like normal adults, tend to be strongly affected by rhyme; poor readers, on the other hand, are significantly less affected. For them, phonetic similarity has relatively little effect on recall (Liberman et al., 1977).

Subsequent experiments have confirmed and extended this result under a variety of conditions: when memory is tested by recognition as well as when it is tested by recall (Byrne & Shea, 1979; Mark, Shankweiler, Liberman, & Fowler, 1977); when sentences or word strings are the stimuli as well as when letter strings are presented (Mann, Liberman, & Shankweiler, 1980); when the items are presented auditorily instead of visually (Shankweiler, Liberman, Mark, Fowler, & Fischer, 1979). In each of these conditions it was found that poor readers are relatively insensitive to the phonetic characteristics of the items. Accordingly, it has been supposed that poor readers have a general problem with the use of a phonetic code, however the material is presented, and not a specific difficulty in deriving a phonetic representation from print (Shankweiler & Liberman, 1976). It would seem, therefore, that one reason for poor readers' deficient performance in short-term memory tasks is their failure to fully exploit phonetic coding.

It remains to be determined what limits full utilization of phonetic codes by poor readers. To what extent does the problem arise in perception with the encoding of stimuli, and to what extent does the problem involve the use of information already represented in phonetic form? Our intent in this study was to investigate whether the poor readers' phonetic-coding deficiency in short-term memory is related to the perceptual process as such.

A study by Rabbitt (1968) gives a way to understand how such a relationship might come about. This study points to a direct connection between stimulus variables that affect perception and those that affect recall. In Rabbitt's experiment, the subjects were required to listen to spoken digits presented with a white noise mask. In one condition the subjects' task was to repeat individual items, in another condition they were tested for recall of strings of items. It was found that noise levels that produced no manifest effect on perception and recall of the individual items significantly impaired recall of the strings. Thus adding noise, and increasing the perceptual difficulty, adversely affected memory even when the individual items could still be identified correctly. The insight we gain from Rabbitt's findings may give us a purchase on the problem of why poor readers typically reveal deficits in verbal short-term memory. Their failure to make full use of phonetic coding in short-term memory may be traceable, as Perfetti and Lesgold have supposed (1979), to a disorder at the level of perceptual processing.

It is well known that severe reading problems often occur in children who show no obvious abnormalities in language development. These poor readers typically do not manifest clinically apparent difficulties in perception of speech. It is conceivable, however, that such children may have subtle deficiencies in speech perception that special testing procedures may bring to light.

One study (Goetzinger, Dirks, & Baer, 1960) hints that in order to discern differences in perceptual skills among good and poor readers it may be necessary to use a quite demanding task. Goetzinger et al. reported no difference between reading groups for a list of well-articulated words but a significant difference in favor of the good readers on a list of rapidly, and somewhat indistinctly, articulated items. Although the study does not permit a direct comparison to be made (different words occurred in the two test lists), the results suggest that discrepancies in speech perception abilities may have been present for good and poor readers that would be detected on a sufficiently difficult task.

Although relevant data are scarce, there is reason to suggest that the characteristic differences so often observed between good and poor readers on memory tasks might be associated with differences in speech perception. Our purpose in the research we present here was to examine this possibility. Accordingly, good and poor readers were tested on a memory task in which the effects of phonetic coding are known to be discernible. Using the procedure of Liberman et al. (1977), we compared performance on recall of phonetically similar (rhyming) and phonetically dissimilar (nonrhyming) sequences of monosyllabic words in good and poor readers. It was expected that, as in previous experiments, good readers, in contrast to poor, would find recall of the rhyming sequences more difficult than the nonrhyming sequences, reflecting more efficient use of a phonetic code. We then addressed the question of whether the reading group differences on memory tasks are related to speech perception abilities. The subjects were tested on a speech perception task requiring repetition of monosyllabic words. The items selected included high and low frequency words phonetically balanced to permit phonetic analysis of errors and examination of error location within the syllable. The stimuli were presented under two conditions, with and without masking noise, in order to vary the difficulty of the task. In addition, a test of perception of environmental nonspeech sounds was conducted, again with and without noise masks, to enable us to investigate any differences in perceptual performance that exist beyond the speech domain.

METHOD

Subjects

The subjects were third-grade children from a suburban public school in southern Rhode Island. A school reading specialist was asked to select the poorest readers and the good readers from the third-grade classes. The children were given the Word Attack and Word Recognition subtests of the Woodcock Reading Mastery Tests, Form A (Woodcock, 1973), and a test of receptive vocabulary, the Peabody Picture Vocabulary Test (PPVT; Dunn, 1965). On the basis of scores obtained on the Woodcock test, two groups were formed that were non-overlapping in reading level.

Eight children were eliminated because their inconsistent scores on the two Woodcock sub-tests made them difficult to classify as good or poor readers. Three additional selection criteria were employed to determine eligibility for participation in the experiments. First, in order to restrict the range of vocabulary skills, only those children were selected whose PPVT IQ score fell between 90 and 120. An additional five children failed to meet this requirement. Second, in view of the evidence that the speech perception skills of children continue to develop during elementary school years (Finkenbinder, 1973; Goldman, Fristoe, & Woodcock, 1970; Schwartz & Goldman, 1974; Thompson, 1963), subjects were selected whose ages fell within a limited range (96 to 108 months). The age requirement excluded five more potential subjects. And third, the remaining children were screened for hearing loss. The right and left ears were presented with tones at 500 Hz (25 db), 1000 Hz (20 dB), 2000 Hz (20 dB), 4000 Hz (20dB), and 8000 Hz (20 dB), using a standard audiometer. Seven children failed the hearing screening.

Thirty children met all the requirements for participation in the study. Table 1 summarizes the characteristics of the good and poor reader groups. The 15 children who qualified as good readers were well ahead of third grade reading skills with a mean reading grade level of 5.88. The 15 children labelled poor readers averaged slightly more than one-half year below their expected level (with a mean reading grade level of 2.76).

Table 1

Means for Third-Grade Children Grouped According to Reading Achievement

Group		Age	IQ[a]	Reading Grade[b]
Good	15	8 yr. 5 mo.	106.8	5.88
Poor	15	8 yr. 6 mo.	102.5	2.76

[a] Peabody Picture Vocabulary Test

[b] From the average of the reading grade scores obtained on the Word Attack and Word Recognition subtests of the Woodcock Reading Mastery Tests, Form A.

The ages of the good (mean = 8 yr. 5 mo.) and poor readers (mean = 8 yr. 6 mo.) did not differ significantly. Nor were the IQ scores as assessed by the PPVT significantly different. The mean IQ score for the good readers was 106.8, for the poor readers 102.5.

Procedure

Each child was tested individually for three sessions. The first session included the screening procedure, the speech perception noise-masked condition and one half (set A as explained below) of the memory experiment. The second session, occurring at least a week later, consisted of the speech perception unmasked condition and the other half (set B) of the memory experiment. The third session, approximately two months after the first, was devoted to the environmental-sounds experiment.

The experiments were conducted in a quiet room. The tape-recorded material for the memory, speech perception, and environmental sounds tasks was played to subjects over earphones. The subjects' responses were recorded on audiotape. Transcriptions of the subjects' responses were also made during the testing session. The tapes were played back within an hour of the experimental session in order to corroborate the transcription and to allow any necessary corrections.

EXPERIMENT 1: Susceptibility of Good and Poor Readers to Phonetic Confusions in Short-term Memory

The first experiment employed a short-term memory task with rhyming and nonrhyming word strings. Our aim was to confirm previous evidence that poor readers make less effective use of phonetic coding in short-term memory than do good readers.

Stimuli

Twenty strings of five monosyllabic words were created, ten rhyming and ten nonrhyming. A single list of 50 common nouns was used as the word source for the rhyming and nonrhyming tests. Thus word frequency, phonetic structure, and word length were strictly controlled for the two conditions. The five words in each rhyming string had the same vowel and the same final consonant if any. The five words in each nonrhyming string all had different vowels and final consonant.

The twenty strings were recorded on magnetic tape in two sets (A and B) of ten lists read by a phonetically-trained male speaker. Each set comprised an alternating presentation of rhyming and nonrhyming strings. Within each string the items were spoken with a neutral prosody at the rate of one per second. The two sets are presented in Table 2.

Procedure

Each subject heard set A during the first session and set B during the second. On both occasions the same procedure was followed.

Table 2

Experiment 1: Word Lists

Set A

1.	chain	train	brain	rain	pain
2.	cat	fly	score	meat	scale
3.	pair	air	hair	chair	bear
4.	roar	wheat	fat	tail	sky
5.	state	plate	weight	gate	fate
6.	tie	hat	nail	floor	sheet
7.	cell	shell	well	bell	spell
8.	mail	pie	store	cap	feet
9.	bee	tree	knee	tea	key
10.	treat	door	eye	sail	map

Set B

1.	bell	state	knee	pain	chair
2.	fly	pie	tie	eye	sky
3.	bee	cell	train	air	plate
4.	cat	hat	fat	map	cap
5.	gate	brain	pair	tea	well
6.	tail	scale	mail	sail	nail
7.	bear	key	weight	shell	chain
8.	score	roar	door	floor	store
9.	rain	hair	spell	fate	tree
10.	meat	wheat	sheet	feet	treat

The child was told that a list of words would be played and that the task was to repeat the list in the order given. After practicing with two lists read by the experimenter, the subject then heard the pre-recorded set of ten five-item word strings.

Results and Discussion

First, an analysis was made of the correct responses in terms of item recall and serial order. Secondly, the errors were analyzed qualitatively in relation to phonetic structure of the stimulus words.

Analysis of Correct Responses

The subjects' responses were scored in two ways. In the first procedure, a response was considered correct only if the item was accurately reported and if it was assigned to the appropriate serial position. The second procedure ignored serial position and counted as correct all responses of words that had occurred in the given string, regardless of order of report.

The error data for each scoring procedure (summarized in Table 3) were subjected to analysis of variance. We examine first the results from the more strict scoring procedure. In agreement with earlier studies (Naidoo, 1970; Miles & Miles, 1977; Shankweiler et al., 1979; Mann et al., 1980) the overall accuracy of recall was greater for good readers, $F(1,28) = 5.6$, $p = .025$. There was as expected, a significant effect of list type, $F(1,28) = 44.2$, $p < .001$. And, as predicted, the good readers made fewer errors on the nonrhyming word sequences than on the rhyming. The poor readers also showed an effect, though a smaller one, of phonetic similarity. Thus, while we obtained significant effects of reader group and of list type that conformed to the pattern of earlier studies (Shankweiler et al., 1979; Mann et al., 1980), the interaction between reading group and list type did not reach significance, $F(1,28) = 2.9$, $p = .098$.

Table 3

Experiment 1: Mean Number Correct Summed Over Serial Positions[2] for Strict Order Scoring and for Order Free Scoring

	Order Correct Scoring		Order Free Scoring	
	Rhyme	Non-Rhyme	Rhyme	Non-Rhyme
Good	15.8	28.0	32.7	35.5
Poor	12.2	19.4	31.7	29.5
difference	3.6	8.6	1.0	6.0

[2] Maximum = 50

Table 4

Experiment 1: Analysis of Incorrect Responses

Reading Group	Condition	incorrect responses / total responses	Composition of incorrect responses: The percentage with:[a] initial consonant	vowel	final consonant	initial consonant and vowel from same word	vowel and final consonant from same word
Good	Rhyme	15	66	99	98	0	79
	Nonrhyme	17	81	97	97	19	68
Poor	Rhyme	16	59	99	97	2	75
	Nonrhyme	26	79	98	92	16	73

---

[a] i.e., the percentage with phonetic information that was available in the two strings.

Evidence that the two reading groups differed in the recall strategies they employed emerges when the data were re-examined after applying the more lenient scoring procedure. As in other studies utilizing lists of high intralist similarity, item information suffers less than order information. So for both groups the order-free recall scores are markedly higher, particularly for the rhyming strings. Overall, the performance level of the two reading groups was not significantly different, $F(1,28) = 3.6$, $p = .071$, nor was there a main effect of rhyme, $F(1,28) = .1$, $p > .500$. In Table 3 we can see, however, that while the scores for the two groups were very close in the rhyming condition, they were dissociated on the nonrhyming sequences. Thus, we find a significant interaction between reading group and list type, $F(1,28) = 6.7$, $p = .016$). The good readers showed improved performance in the nonrhyming condition, $F(1,28) = 4.2$, $p = .05$, where an efficient phonetic strategy can operate to advantage. The poor readres, in contrast, did not improve on the nonrhyming sequences, $F(1,28) = 2.6$, $p < .20$; indeed they tended to do worse.

The memory experiment undertaken here was intended mainly as a replication. In previous research, good readers evidenced generally superior recall but were relatively more penalized by phonetic similarity within a list than were poor readers. The present study does generally conform to this picture, though here the differences between the groups were somewhat less marked, perhaps because the subjects were a year older than those in the earlier research. At present, the appropriate studies to examine developmental changes in use of a phonetic strategy have not been done. If poor readers are employing a non-phonetic strategy, as has been suggested (see Byrne & Shea, 1979), we might expect their use of this strategy to diminish with increasing age (Conrad, 1972).

Qualitative Analysis of Errors

The construction of the present experiment, using words as stimuli rather than letters, permits a closer inspection of the nature of the difficulty poor readers have in preserving order information. In analyzing the response sequences, it became apparent that the recall problems of poor readers apply not only to the order of the stimuli in a string but also to the retention of phonemic sequences within individual words. The subjects' response sequences (for both good and poor readers) included items that had not occurred in the strings. These errors were often obvious recombinations of phonetic components that had been present in the presented sequence (e.g., for the target items *train* and *plate* several subjects reported *trait* and *plane*). Such errors of transposition have previously been reported in memory experiments with adults (Drewnowski, 1980; Ellis, 1980). We undertook to analyze the phonetic errors in the present experiment to determine how often the incorrect responses could be accounted for as transposed phonetic segments from adjacent items. In this analysis, the given string and the previous sequence were considered as the available source of phonetic information.

The data base for determining whether errors of transposition were present was the 451 phonetic errors obtained from all 30 subjects. Seven of these errors were whole words from previous lists and were disregarded. An additional seven were discounted because they were phonetically unrelated to any item in either word list. The phonetic composition of the remaining 437 responses could, for the most part, be accounted for in terms of the phonetic units present in the particular string and the preceding string. In Table 4 we present a breakdown of the transposition errors. Good and poor readers' transposition errors were very similar in pattern. When a phonetic unit was transposed, it was recombined in the same syllable position in which it had originally occurred. Most commonly, vowel and final consonant (or consonant cluster) were preserved as a unit with a substituted initial consonant (or consonant cluster). (Table 5 lists a representative sample of the observed error responses.) This error pattern suggests that phonetic segments are not equally free to dissociate and recombine in memory. If they did operate as independent units on recombination, there would be no reason to expect greater cohesion between the vowel and the final consonant than between the initial consonant and the vowel.

Table 5

Experiment 1: Examples of Transposition Errors

Presented Items	Responses
roar + fat	rat
bear + shell	bell
score + cat	scat
knee + state	neat
chair + pain	chain
hair + spell	hell
spell + fate	spate
pie + feat	peat
tea + brain	tain

To ascertain whether the incidence of transposition errors differentiates the reading groups, an analysis of variance was carried out on the proportion of transposition errors to correct responses for the rhyming and nonrhyming conditions. The overall proportion of transposition errors to correct responses did not differ significantly for the two reading groups, $F(1,28) = 1.8$, $p = .194$. However, while both groups produced a higher proportion of transposed responses in the nonrhyming condition, the difference was more pronounced for the poor readers. These effects are manifested by a significant effect of list type, $F(1,28) = 10.4$, $p = .004$, and by a significant interaction between list type and reading group, $F(1,28) = 4.9$, $p = .036$. Thus it seems that the greater difficulty poor readers have in retaining the order of words in the nonrhyming sequences may be compounded by a problem with the preservation of order information within a word. In the case of the rhyming strings, of course, subjects may well produce transposed responses that would be undetectable. This may account for the better performance of the poor readers in the order-free scoring of rhyming words.

The present study confirms earlier reports that poor readers recall fewer items than good readers and that they are less affected by phonetic similarity within a list than are good readers (Liberman et al., 1977; Mann et al., 1980; Mark et al., 1977; Shankweiler et al., 1979). In this study the result of the phonetic error analysis allows us to extend our understanding of poor readers' performance on memory tasks. It indicates first of all that the poor readers definitely obtained the phonetic information in the stimuli. However, the greater incidence of transposition errors by poor readers (in the nonrhyming condition) also points to inferior retention of the correct combinations of phonetic sequences specifying the individual items. This finding is consistent with other indications (Katz, Shankweiler, & Liberman, in press) tnat poor readers encounter difficulty in preserving serial order information in linguistic tasks. It further suggests that the problem extends to the ordering of segments within the syllable.

EXPERIMENT 2: Speech Perception in Good and Poor Readers

We now turn to the second question: the speech perception abilities of the good and poor readers. The aim of Experiment 2 was to investigate whether the language deficits of the poor reader are evident in phonetic perception as well as in short-term memory.

Stimuli

The perception test consisted of 48 words especially chosen to control for syllable pattern, phonetic composition, and word frequency. There were 12 words for each of the following syllabic patterns: CVC (consonant-vowel-consonant), CCVC, CCVCC, and CVCC. Within each syllable pattern, half of the words selected were judged to have high frequency of occurrence in children's literature and half had low frequency (Carroll, Davies, & Richman, 1971). The frequency values were validated with a second word frequency source (Thorndike & Lorge, 1944).

In order to permit a clearcut analysis of phonetic errors and of errors of position (i.e., initial, medial and final word position) words were chosen to provide a systematic phonetic set. Twenty words began with stop consonants (/b/, /d/, /[illegible]/, /p/, /t/, /k/) and twenty words began with fricatives or affricates (/tʃ/, /s/, /f/, /ʃ/, /dʒ/, /v/). For each of the above phonetic categories half of the occurrences were in high frequency words and half were in low frequency words. Of the remaining eight items, four began with nasal consonants (/m/, /n/) and four with liquids (/r/, /l/). The same distribution of phonetic elements occurred in word final position.[1]

The occurrences of segments in medial position were not controlled except in one respect: every syllabic pattern that occurred in a high frequency word was matched in a low frequency word (e.g., front [high frequency] and flint [low frequency] were matched in syllabic pattern: each consisted of the sequence: fricative, liquid, vowel, nasal consonant, stop consonant). The word list is presented in Table 6.

Table 6

Experiment 2: Speech Stimuli

High Frequency Words	Low Frequency Words
door	bale
team	din
road	lobe
knife	mash
chief	chef
job	fig
grain	tram
breath	grouse
crowd	crag
sleep	slag
scale	spire
speech	skiff
front	flint
plant	clamp
friend	frond
clouds	glades
blocks	drapes
planes	prunes
bank	kink
chance	finch
list	rasp
month	nymph
child	vault
ships	shacks

The words were recorded by a phonetically-trained male speaker, each being produced as the final word of a meaningful sentence. The sentences were subsequently digitized at 10,000 samples/sec and each stimulus word was excised from the rest of the sentence, using the Haskins WENDY waveform editing system (Szubowicz, Note 1). The words were then arranged into a fixed random sequence and recorded onto magnetic tape. When the stimuli were replayed, a comfortable listening level was selected, approximately 78 dB SPL.

The noise-masked condition was then constructed by following the method described by Schroeder (1968). The technique involves computing the masking noise signal directly from the digitized speech sample to be masked. Each speech sample of the digitized waveform of a stimulus word is multipled by another, randomly chosen with equal probability. The waveform that results from this manipulation preserves the time-varying amplitude characteristics of the speech signal while having a flat long-term frequency spectrum. Thus it is referred to as an amplitude-match noise signal. Each digitized word and its amplitude matched noise signal were added linearly to yield a 0 dB S/N ratio. The words in noise were subsequently arranged into a fixed random order and recorded on magnetic tape.

Procedure

Each subject listened to the noise-masked words during session 1, and the unmasked words during session 2. The child was told that a list of words would be played (and, in the noise-masked condition, that the words were recorded in some noise). The subjects were instructed to repeat each item clearly immediately after hearing it. The test sequence was preceded by four practice trials.

Results and Discussion

Few words were missed by either the good readers (mean errors = 1.3) or the poor (mean errors = 2.0) in the unmasked condition. As we can see in the left-hand portion of Figure 1, whereas both groups made considerably more errors in the noise-masked condition, the poor readers (mean errors = 20.7) did markedly worse than the good readers (mean errors = 15.1).

These effects were analyzed by a two-way factorial analysis of variance. The between-groups factor, reading achievement, was significant, $F(1,28) = 17.6$, $p < .001$, with good readers misreporting fewer words than poor readers. In addition, there was a significant main effect of noise, $F(1,28) = 687.4$, $p < .001$. From previous perception research with adults (e.g., Licklider & Miller, 1951), the detrimental effect of masking noise on intelligibility is well known. What is new, from our point of view, was the finding that there were notable differences in the magnitude of the effect of noise on perception for the two reading groups. A significant interaction between the effect of masking and reading group was obtained, $F(1,28) = 15.8$, $p < .001$. When the stimuli were presented clearly in the unmasked condition, all the subjects reported the stimulus items accurately. The addition of noise, however, made it significantly more difficult for the poor readers to perceive the stimuli than for the good readers to do so. Thus it seems that the speech perception skills of poor readers are less effective than those of good readers but that this difference is observable only when they are required to respond to degraded stimuli.

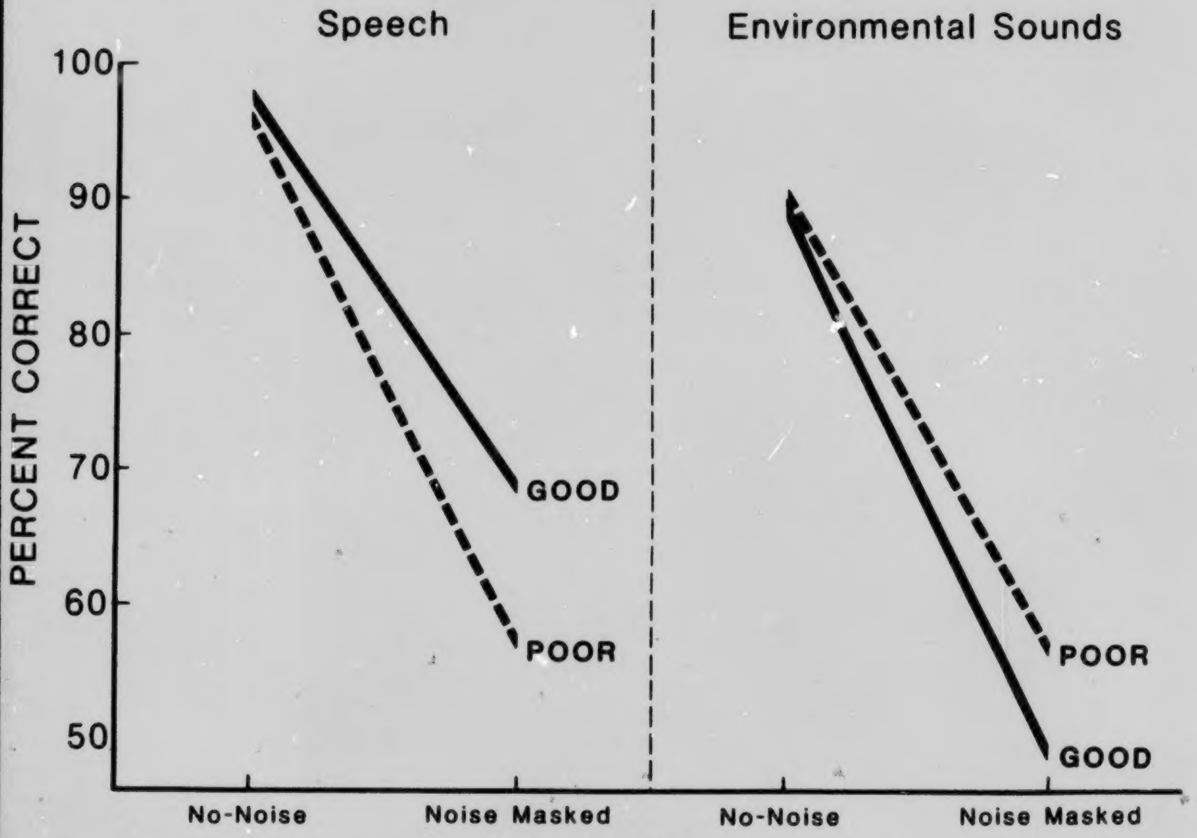


Figure 1. Performance of good and poor readers on the speech perception task (Experiment 2) and the environmental sounds task (Experiment 3), plotted in mean percent correct.

Words of high and low frequency of occurrence were employed in the experiment as a means of examining whether differences between the groups in perceptibility of the items were attributable to differences in vocabulary skills. In Figure 2 we can see the performance of the two reading groups on the high and low frequency items. While the variable of word frequency had a large effect on the perceptibility of a word, $F(1,28) = 155.0$, $p < .001$, there was no interaction between the word frequency variable and reading group, $F(1,28) = .015$, $p > .500$. The poorer performance of the poor readers cannot, therefore, be attributed to possible differences in word knowledge. Instead, it points to a problem in perception of speech.

Thus far we have examined the results by viewing each response either as being totally correct or as an error. In order to determine where the perceptual mistakes were occurring, it is useful to examine the nature of the errors as was done by Shankweiler and Liberman (1972). Accordingly, each stimulus was broken into three segments: the initial cluster, the medial vowel and the final cluster. A given error response could deviate from the target stimulus at one, two, or all three word positions. The error data for this analysis are summarized in Table 7. For both reading groups, the greatest number of errors occurred in the initial portion of the word, the final position was second in error rate, and very few errors were made on the vowel in medial position. This position effect was significant, $F(2,55) = 169.2$, $p < .001$, with no difference in error pattern between the good and poor readers. The lack of an interaction between position effect and reading group suggests that the basis for the error pattern was the same for both good and poor readers. We will briefly digress to consider what these factors might have been.

Table 7

Experiment 2: Speech-in-Noise: Error Location Within the Stimuli[2]

	Initial	Medial	Final	
Good	11.27	2.2	7.07	Mean number of errors on 48 trials
Poor	14.67	3.7	8.93	

[2] Error position not exclusive

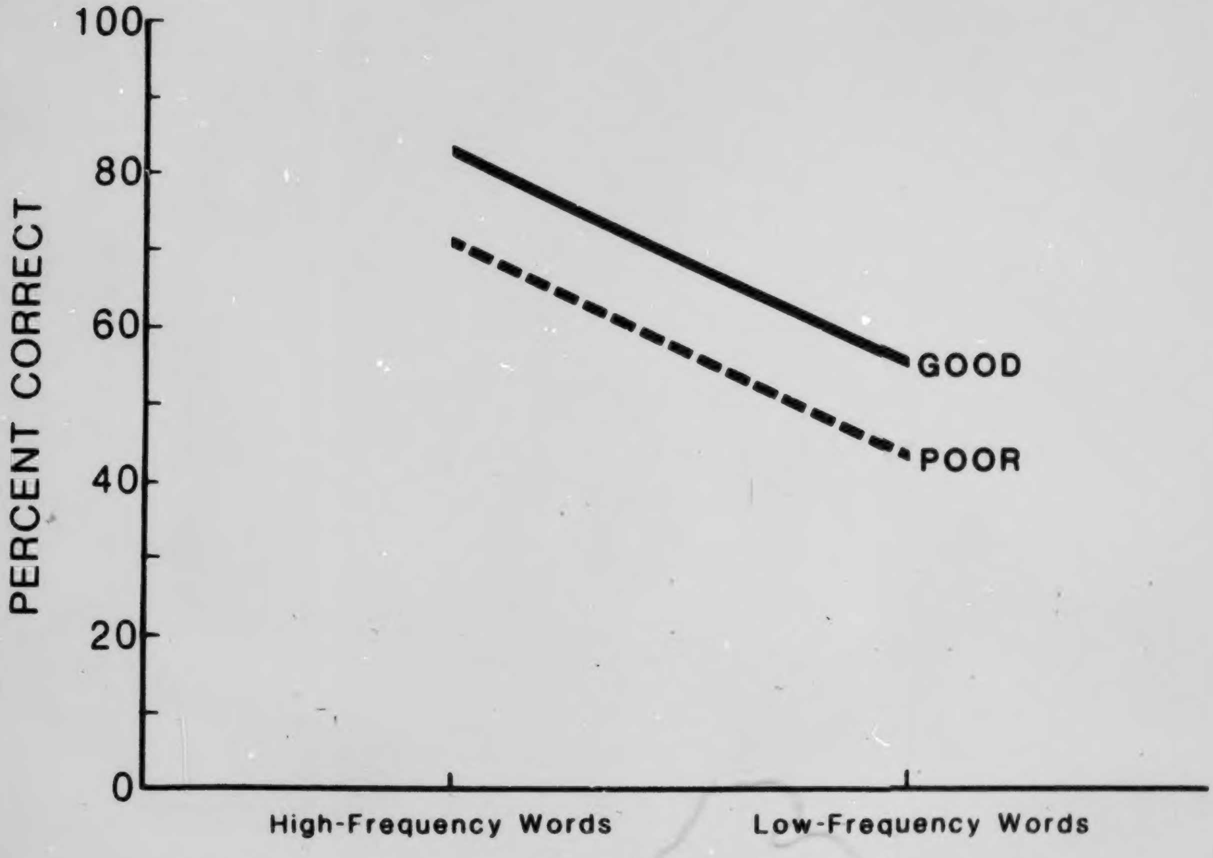


Figure 2. Speech perception (Experiment 2): Mean percent correct on the noise-masked condition replotted as a function of word frequency.

The uneven distribution of errors across the three word positions seems to correspond with the relative acoustic saliency of the segments. The vowel in acoustic terms is more intense than consonants and is longer in duration. It is therefore not surprising to observe superior identification of vowels on a listening task. Our finding that the initial consonant (or consonant cluster) is misheard more often than the final consonant (or consonant cluster) parallels research with CV and VC syllables (see Ohde & Sharf, 1977, for a major paper in this area; and Ohde & Sharf, 1981, and Pols & Schouten, 1981, for recent discussions of those findings), and again seems to be related to the acoustic characteristics of the segments. The results of research on the speech cues suggest that the consonant in final position is more clearly represented in the acoustic signal than is the initial consonant. Syllable final formants have been observed to have transitions of greater duration (except following the vowels /e/ and /i/) (Lehiste & Peterson, 1961) and greater frequency change (Broad & Fertig, 1970) than have initial transitions. Further, the vowel nucleus of the syllable has been found to provide a variety of cues that may aid in identification of final segments. Peterson and Lehiste, 1960, observed vowel lengthening accompanying voiced final fricatives and voiced final consonants, and greater nasalization of vowels preceding nasal consonants than for vowels following nasal consonants. Thus final consonants may be easier to perceive because a greater amount of information specifies their identity.

In view of the position effects obtained here, it seemed appropriate to examine the phonetic composition of errors occurring in initial and final position. For both positions, an adequate sampling was available to compare the relative frequencies of occurrence of errors on stop consonants and fricatives (see Table 8), but not on liquids or nasals. Accordingly, an analysis of variance was carried out on the stop consonant errors and the fricative errors with error position, initial or final, specified. In this analysis our previous findings were again substantiated: good readers made fewer errors, $F(1,28) = 10.0$, $p = .004$; more errors occurred on initial position than on final, $F(1,28) = 51.2$, $p < .001$; and there was no interaction between reading groups and the position effect. A significant difference was obtained between the two phonetic categories examined. More stop consonants were missed than fricatives, $F(1,28) = 51.1$, $p < .001$ and an interaction between reading group and phonetic category was obtained, $F(1,28) = 5.4$, $p = .03$. The poor readers missed the stop consonants significantly more often than did the good readers. This could be taken as an indication that poor readers have particular difficulty in processing stop consonants. At the present, we are inclined to make the more conservative speculation that, with the particular noise utilized, the stop information in the signals was relatively more obscured than was fricative information. Given that the amplitude characteristics of the word were preserved in the noise signal, an important cue for fricative identity would also be preserved while place information for the stops would be less salient.

Table 8

Experiment 2:
Speech-in-Noise: Analysis of Error Position and Phonetic Category

Relative occurrence of errors of a phonetic category: e.g. the stop consonants missed in initial position/the stop consonants that occurred in the initial cluster.

	Initial Position		Final Position	
	Stops	Fricatives	Stops	Fricatives
Good	25.1	12.7	11.8	12.9
Poor	35.4	17.7	17.7	13.4

In sum, we found that on the unmasked condition the poor readers did as well as the good readers. When the perceptual system was stressed by the addition of noise, the poor readers made significantly more errors in perceiving the stimuli than did the good readers. With these results in hand, we may now consider the question whether the difficulties the poor reader has with reading may stem from a more general problem in auditory perception. If poor readers are generally inferior to good readers on another auditory perception task, where speech processing is not required, a different interpretation of the nature of the poor readers' problem would be necessary than would be appropriate if the problem were specific to speech.

EXPERIMENT 3: Perception of Environmental Sounds With and Without Noise Masks

In the final experiment the subjects listened to a tape of environmental sounds: first with the stimuli in noise, then in quiet.

Stimuli

The stimuli for this experiment were selected and edited from a magnetic tape recording of environmental sounds that had been obtained from the Neuropsychology Laboratory at the University of Victoria (Spreen & Benton, 1969). The source tape had 26 sounds, two of which were excluded for use here because they contained speech. The remaining 24 stimuli, listed in Table 9, included human nonspeech sounds (e.g., coughing), human activities (e.g., knocking on a door), mechanical sounds (e.g., machine-gun fire), animal noises

Table 9

Experiment 3: Environmental Sounds Stimuli

1. Knocking on a door
2. Water running from a faucet
3. Organ - wedding march
4. Phone ringing
5. Whistling
6. Airplane engine
7. Door opening and closing
8. Artillery
9. Car starting up and driving away
10. Dialing a phone
11. Drum
12. Birds
13. Church bell-time
14. Frogs and crickets
15. Piano
16. Dog barking
17. Trumpet fanfare
18. Train whistle
19. Cat meowing
20. Clapping
21. Coughing
22. Baby crying
23. Thunder
24. Typing

(e.g., frog croaks and cricket chirps), and sounds of nature (e.g., thunder). Each sound was digitized on the Haskins Laboratories DDP-224 PCM system and recorded on magnetic tape. One taped sequence, for the unmasked condition, contained the sounds presented in a fixed random order. In constructing the noise-masked sequence, it was not advantageous to use amplitude matched noises as we had done in the case of the speech perception experiment, since the amplitude characteristics of the environmental sounds often provided strong cues to the identity of those sounds. We therefore chose instead to use a broad band (0 to 10 kHz) white-noise signal as the masking stimulus. Pilot work suggested that a 0 dB S/N ratio, as employed in the speech task, did not sufficiently mask the stimuli, but that a -2 dB S/N ratio would be appropriate. A second sequence for the noise-masked condition was recorded with each sound masked by the white noise signal at the -2 dB S/N ratio. The stimuli for the two listening conditions were replayed at a comfortable listening level of approximately 75 dB SPL.

Procedure

Both the noise-masked and the unmasked stimuli were presented in a single session, with all subjects listening to the noise-masked tape first. Prior to the testing the examiner explained that the child would hear two sets of sounds and that in the first set the items were recorded with noise. The child was asked to identify the source of each sound immediately after hearing it, providing as much detail as possible. Three practice trials were conducted, without noise, to familiarize the subject with describing nonspeech sounds.

Results and Discussion

The subjects' responses were compiled into a single list. Before scoring, all the responses to each sound were evaluated. A point system was devised ranging from 0 to 3. A score of zero was assigned if the response bore no relation to the stimulus; three was awarded if a fully specific identification had been provided. For the intermediate scores, a score of one was given if the response reflected the nature of the sound though wrong in detail (e.g., for coughing, if the S responded 'talking' or 'laughing' that person had correctly determined that a human vocal tract was the source); two was assigned if the response was not inaccurate but somewhat unspecific (e.g., for an organ playing the wedding march, the response 'music'). Responses distributed themselves somewhat unevenly: for some of the stimuli not all four of the scoring categories were assigned. The scoring was reviewed by a colleague who did not know which responses came from good readers and which from poor ones. Discrepancies in numerical assignment by the two scorers occurred for two responses and these were resolved by joint discussion of the two cases. The subjects' answer sheets were then scored and tabulated. The mean error score in the unmasked condition was 6.7 for the poor readers and 7.6 for the good readers (maximum = 72). In the noise-masked condition the mean error scores were 31.4 for the poor readers and 36.9 for the good readers. These performance levels are displayed in the right-hand portion of Figure 1.

As in the speech perception experiment, few errors were made by either reading group in the unmasked condition. With the addition of masking noise, performance for both groups was markedly reduced. The analysis of variance revealed a main effect of noise, $F(1,28) = 510.9$, $p < .001$, and a main effect of reading group, $F(1,28) = 4.7$, $p = .04$. We note that the poor readers performed better than the good readers on the nonspeech task. However, if age and IQ are controlled, the difference did not reach significance, $F(1,26) = 3.6$, $p = .071$.[2] Given the equality of the performance of the poor readers with that of the good readers on this nonspeech auditory task, we can rule out inattention as the explanation for their inferior performance on the noise-masked speech perception task. The results of this control experiment further suggest that the difficulty the poor readers manifested in perceiving speech in noise is not the consequence of generally deficient auditory perceptual ability, but rather is related specifically to the processing requirements for speech.

DISCUSSION

Earlier work has demonstrated that children who are poor readers have short-term memory deficits in situations where the stimuli lend themselves to phonetic coding. The present experiments were intended to investigate the basis of this deficit, by asking whether the language processing problems of poor readers may extend to the area of phonetic perception. Third-grade school children selected for reading ability were first tested on serial recall of word strings, a task that previously had been found to differentiate good and poor readers (Mann et al., 1980). As before, the poor readers made more errors than the good readers. The results are consistent with the hypothesis (Liberman et al., 1977; Shankweiler et al., 1979) that a failure to use phonetic coding efficiently leads to the poor reader's deficiency in short-term memory for labelable stimuli.

In order to investigate the origin of this memory coding problem, the subjects were further tested on two tasks. One of these employed spoken words and the other, nonspeech environmental sounds. Each task was presented under two conditions: one with a favorable signal-to-noise ratio and one with masking noise. The results indicated a deficit for the poor reader group that was specific to speech stimuli and occurred only in the noise-masked condition. Significantly more errors were made by the poor readers than the good readers when listening to speech in noise; the groups did not differ, however, in the perception of nonspeech environmental sounds, whether noise-masked or not. This pattern of results suggests that the poor readers could process the speech signal adequately, as expected, but they required a higher quality signal for error-free performance than the good readers. The absence of differences between the reading groups on the control experiment with environmental sounds suggests that the poor readers' problem is not manifest on just any auditory task in which the stimuli are noisy, but is instead more selective. The joint outcome of these perception studies suggests that poor readers require more complete stimulus information than good readers in order to apprehend the phonetic shape of spoken words.

The present experiment has demonstrated associated deficits on the same group of poor readers: inferior performance on serial recall and inferior performance on a stringent test of speech perception. We now turn to consider how these two deficits might be related. First, we have noted that poor readers show weak effects of phonetic similarity in recall tasks, a fact that has been taken as evidence that they make inefficient use of phonetic coding in short-term memory. In the memory experiment of the present study, the analysis of the error responses provides direct evidence that the poor readers were using a phonetic code to retain material in short-term memory, though, of course, less effectively than the good readers. The errors that occurred were rarely semantically related to the target items, which might have indicated use of an alternative coding strategy; instead, they consisted of transpositions of phonetic segments from adjacent syllables. Such an error pattern as we obtained seems possible only if the subjects were indeed using a phonetic coding strategy. Thus, it is apparent that whereas both good and poor readers were phonetically coding the stimuli, the poor readers were more apt to exchange segments across word boundaries and they experienced greater difficulty in retaining the order of words within each word string.

Thus the suggestion that poor readers have greater difficulty in correctly retaining phonetic representations is corroborated by the pattern of their errors on the serial recall task. In the word perception task, we obtained evidence that poor readers also experience greater difficulty perceiving the phonetic form. On the contrary, analysis of errors in word perception showed that good and poor readers did not differ in the effect of word frequency on item identifiability. Therefore, the greater susceptibility of the poor readers to errors of identification apparently does not arise from differences between good and poor readers in vocabulary level. The problem thus appears to be not in dealing with the linguistic content of the stimulus items, but rather with the form. In perception as well as in recall of linguistic items, the poor readers' problems would seem to stem from failure to adequately internalize certain formal properties of language: in these instances, properties relating to the phonetic pattern.

We may speculate therefore that the problems of poor readers, evident on both the memory task and the perceptual task, arise at least in part from a common cause. In this connection, it may be relevant to recall the finding by Rabbitt (1968), to which we have referred, in which there was shown to be a relationship between recall performance and the stimulus factors that affect perceptual clarity. When adult subjects were asked to recall strings of digits, recall of items presented without noise was impeded if subsequent items were presented in noise. Thus, making some items difficult to perceive seems to reduce ability to rehearse the non-noisy items of the string also. We may speculate, by extension, that poor readers' recall suffers in part from the difficulties they incur in perceptual processing.

Thus one may surmise from our results that the recall performance of poor readers for words presented auditorily suffers as a result of faulty phonetic coding of the stimuli. Moreover, we suppose that this difficulty may arise whenever a phonetic representation is formed irrespective of the sensory modality of the signal. We base this conjecture on the outcome of earlier findings (Liberman et al., 1977; Shankweiler et al., 1979) which have shown that the failure of poor readers to make full use of phonetic coding in recall occurs both with auditory presentation and with visual presentation of the stimulus items. These parallel findings for presentation of stimuli by ear or by eye led us to suppose that poor readers' problems in memory coding are of a linguistic nature.

It is noteworthy that other investigators who have employed similar criteria for subject selection, but who have used very different experimental approaches to that adopted in the present study, have reached a similar conclusion. Using the memory scan procedure of Sternberg (1966), Katz and Wicklund (1971) have found slower encoding times for poor readers than for good readers with visually-presented word-strings. If we are correct in supposing that the memory deficit in poor readers at least in part has its origin in phonetic perception, it should be possible to demonstrate differences in a variety of situations in the facility and accuracy with which good and poor readers process linguistically codable material that is presented either visually or auditorily.

REFERENCE NOTES

1. Szubowicz, L. *A tutorial guide to Wendy: The Haskins wave editing and display system.* Haskins Laboratories, 1977.

REFERENCES

Baddeley, A. D. Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology*, 1966, *18*, 362-365.

Broad, D. J., & Fertig, R. H. Formant-frequency trajectories in selected CVC-syllable nuclei. *Journal of the Acoustical Society of America*, 1970, *47*, 1572-1582.

Byrne, B., & Shea, P. Semantic and phonetic memory in beginning readers. *Memory & Cognition*, 1979, *7*, 333-338.

Carroll, J. B., Davies, P., & Richman, B. *Word frequency book*. New York: American Heritage Publishing Company, Inc., 1971.

Conrad, R. Acoustic confusions in immediate memory. *British Journal of Psychology*, 1964, *55*, 75-84.

Conrad, R. Speech and reading. In J. Kavanagh & I. Mattingly (Eds.), *Language by ear and by eye: The relationships between speech and reading*. Cambridge, Mass.: MIT Press, 1972, 205-240.

Drewnowski, A. Memory functions for vowels and consonants: A reinterpretation of acoustic similarity effects. *Journal of Verbal Learning and Verbal Behavior*, 1980, *19*, 176-193.

Dunn, L. M. *Peabody picture vocabulary test*. Circle Pines, Mn.: American Guidance Service, 1965.

Ellis, A. W. Errors in speech and short-term memory: The effects of phonemic similarity and syllable position. *Journal of Verbal Learning and Verbal Behavior*, 1980, *19*, 624-634.

Finkenbinder, R. L. A descriptive study of the Goldman-Fristoe-Woodcock test of auditory discrimination and selected reading variables with primary school children. *The Journal of Special Education*, 1973, *7*, 125-131.

Goetzinger, C., Dirks, D., & Baer, C. J. Auditory discrimination and visual perception in good and poor readers. *Annals of Otology, Rhinology and Laryngology*, 1960, *69*, 121-136.

Goldman, R., Fristoe, M., & Woodcock, R. W. *Goldman-Fristoe-Woodcock test of auditory discrimination*. Circle Pines, Mn.: American Guidance Service, 1970.

Katz, L., & Wicklund, D. A. Word scanning rate for good and poor readers. Journal of Educational Psychology, 1971, 63, 138-140.

Katz, R. B., Shankweiler, D., & Liberman, I. Y. Memory for item order and phonetic recoding in the beginning reader. Journal of Experimental Child Psychology, in press.

Lehiste, I., & Peterson, G. E. Transitions, glides and diphthongs. Journal of the Acoustical Society of America, 1961, 33, 268-277.

Liberman, I. Y., Shankweiler, D., Liberman, A. M., Fowler, C., & Fischer, F. W. Phonetic segmentation and recoding in the beginning reader. In A. S. Reber & D. L. Scarborough (Eds.), Toward a psychology of reading: The proceedings of the CUNY Conferences. Hillsdale, N.J.: Erlbaum, 1977, 207-225.

Licklider, J. C. R., & Miller, G. The perception of speech. In S. S. Stevens (Ed.), Handbook of experimental psychology. New York: Wiley, 1951.

Mann, V. A., Liberman, I. Y., & Shankweiler, D. Children's memory for sentences and word strings in relation to reading ability. Memory & Cognition, 1980, 8, 329-335.

Mark, L. S., Shankweiler, D., Liberman, I. Y., & Fowler, C. A. Phonetic recoding and reading difficulty in beginning readers. Memory & Cognition, 1977, 5, 623-629.

Miles, N. C., & Miles, T. R. Dyslexia as a limitation in the ability to process information. Bulletin of the Orton Society, 1977, 27, 72-81.

Naidoo, S. The assessment of dyslexic children. In A. W. Franklin & S. Naidoo (Eds.), Assessment and teaching of dyslexic children. London: Invalid Children's Aid Association, 1970.

Ohde, R. N., & Sharf, D. J. Order effect of acoustic segments of VC and CV syllables on stop and vowel identification. Journal of Speech and Hearing Research, 1977, 20, 543-554.

Ohde, R. N., & Sharf, D. J. Stop identification from vocalic transition plus vowel segments of CV and VC syllables: A follow-up study. Journal of the Acoustical Society of America, 1981, 69, 297-300.

Perfetti, C. A., & Lesgold, A. M. Discourse comprehension and sources of individual differences. In M. Just & P. Carpenter (Eds.), Cognitive processes in comprehension. Hillsdale, N.J.: Erlbaum, 1979.

Peterson, G. E., & Lehiste, I. Duration of syllable nuclei in English. Journal of the Acoustical Society of America, 1960, 32, 693-703.

Pols, L. C. W., & Schouten, M. E. H. Identification of deleted plosives: The effect of adding noise or applying a time window (A reply to Ohde and Sharf). Journal of the Acoustical Society of America, 1981, 69, 301-303.

Rabbitt, P. M. A. Channel-capacity, intelligibility and immediate memory. Quarterly Journal of Experimental Psychology, 1968, 20, 241-248.

Schroeder, M. R. Reference signal for signal quality studies. Journal of the Acoustical Society of America, 1968, 44, 1735-136.

Schwartz, A., & Goldman, R. Variables influencing performance on speech-sound discrimination tests. Journal of Speech and Hearing Research, 1974, 17, 25-32.

Shankweiler, D., & Liberman, I. Y. Misreading: A search for causes. In J. F. Kavanagh & I. G. Mattingly (Eds.), Language by ear and by eye: The relationship between speech and reading. Cambridge, Mass.: The MIT Press, 1972, 293-317.

Shankweiler, D., & Liberman, I. Y. Exploring the relations between reading and speech. In R. M. Knights & D. K. Bakker (Eds.), Neuropsychology of learning disorders: Theoretical approaches. Baltimore: University Park Press, 1976, 297-313.

Shankweiler, D., Liberman, I. Y., Mark, L. S., Fowler, C. A., & Fischer, F. W. The speech code and learning to read. Journal of Experimental Psychology: Human Learning and Memory, 1979, 5, 531-545.

Spreen, O., & Benton, A. L. Sound recognition test. Victoria, B.C.: Neuropsychology Laboratory, University of Victoria, 1969.

Sternberg, S. High-speed scanning in human memory. Science, 1966, 153, 652-654.

Thompson, B. A longitudinal study of auditory discrimination. Journal of Educational Research, 1963, 56, 376-378.

Thorndike, E. L., & Lorge, I. The teacher's word book of 30,000 words. New York: College Press, 1944.

Woodcock, R. W. Woodcock Reading Mastery Tests. Circle Pines, Mn.: American Guidance Services, Inc., 1973.

FOOTNOTES

[1] In word final position the fricative and affricate set was slightly different, consisting of /f/, /s/, /tʃ/, /ʃ/, /θ/ and /z/.

[2] In Experiments 1 and 2, the data were likewise reanalyzed controlling for age and IQ. In these experiments, the significance of the differences between reading groups was not reduced when age and IQ were controlled.

THE USE OF ORTHOGRAPHIC STRUCTURE BY DEAF ADULTS: RECOGNITION OF FINGERSPELLED LETTERS

Vicki L. Hanson

Abstract. Deaf adults' knowledge of English word structure was tested in a task requiring letter report for fingerspelled words, pseudowords, and nonwords. Deaf subjects, like hearing subjects, were sensitive to orthographic structure as indicated by accuracy of letter report: Letters of words were reported most accurately, while letters of pseudowords were reported more accurately than letters of nonwords. Analysis of the incorrect letter reports for correctly recognized words revealed that deaf subjects tended to produce orthographically regular responses. However, in contrast to the reports of hearing subjects, the responses of deaf subjects did not tend to be phonetically consistent with the presented word. These results provide clear evidence that deaf adults are able to abstract principles of English orthography, although the phonetically inconsistent letter reports suggest that the spelling process for deaf persons may be fundamentally different from that for hearing persons.

The present research examines the use of orthographic structure by prelingually and profoundly deaf adults. The orthography of English reflects the phonological structure of the spoken language. As a result, segments of

---

*An earlier version of this paper was presented at the meeting of the American Psychological Association, Los Angeles, August, 1981.

Acknowledgment. I am grateful to many people for their help on this project. First, I would like to thank Carol Padden for making the stimulus tape. I would also like to thank the people who made arrangements for facilities and subjects for the reported experiment and for pilot work: Nancy Fishbein, Nancy Frishberg, Peg Hlibok, Gary Scharff, Dennis Schemenauer, and Marie Taccogna. The cooperation of the following organizations and institutions is gratefully acknowledged: National Center on Deafness at California State University, Northridge; New York University; New York Society for the Deaf; Connecticut Commission for the Deaf and Hearing Impaired; and the Linguistics Research Laboratory at Gallaudet College. John Richards graciously helped with portions of the data analysis. The handshapes in Figure 1 were drawn by John Conti. Portions of this research were conducted while the author was a postdoctoral fellow at The Salk Institute for Biological Studies. The work was supported by the National Institute of Education Grant #NIE-G-80-0178 and by NINCDS Research Service Award #NS06109 and NICHD Grant HD-01994.

[HASKINS LABORATORIES: Status Report on Speech Research SR-69 (1982)]

the written language map onto segments of the spoken language. The question here is whether deaf adults, in the absence of normal speech input, are able to abstract the regularities of English orthographic structure.

Ability to use the regularities of the orthography is an important component both in word recognition and in spelling. Research on word recognition with normally-hearing adults has found that there is an advantage in letter recognition for orthographically regular nonsense words (pseudowords) over orthographically irregular nonwords (Aderman & Smith, 1971; Baron & Thurston, 1973; Carr, Davidson, & Hawkins, 1978) and an advantage in letter recall for these regular over irregular nonsense words (Gibson, Pick, Osser, & Hammond, 1962).

In spelling, the ability to access and exploit the orthographic regularities of English is a factor determining spelling success. While accurate spelling of words can result from rote memorization or from visual recognition of the correct spelling from a collection of possible spellings (Simon & Simon, 1973; Tenney, 1980), these strategies ignore the systematic aspects of English orthography (Chomsky, 1970; Klima, 1972; Venezky, 1970). Recent work by Fischer (1980) has shown that good spellers have greater ability to exploit these regularities of the orthography than do poor spellers.

To date, little work has been concerned with the question of use of orthographic structure by deaf individuals. One study that has been directed at this issue is that of Gibson, Shurcliff, and Yonas (1970). Testing for recall of tachistiscopically presented pseudowords and nonwords, they found that deaf adults, like hearing adults, correctly recalled more of the orthographically regular than of the orthographically irregular letter strings. Similar findings were obtained by Doehring & Rosenstein (1960) in an experiment with deaf children (ages 9-16 years). They found better recall of CVC trigrams (pseudowords) than of CCC trigrams (nonwords). These findings led Gibson et al. (1970) to conclude that "The redundancy contributed by invariant mapping of speech sounds may well make it easier for the hearing child to pick up the common spelling patterns and regularities as he learns to read, but clearly it can be done without this" (p. 71).

The present research examined their conclusion. The ability of deaf adults to use orthographic structure in word recognition and in reporting the letters of words was tested. Deaf subjects in this research were all congenitally and profoundly deaf adults. These persons are unable to acquire knowledge of speech by normal means. Since the orthography of English reflects the structure of the spoken language, these deaf adults may be expected to be less able than hearing adults to acquire knowledge of this structure and to use it. If, however, as suggested by Gibson et al. (1970), ability to acquire knowledge of orthographic structure does not depend on availability of normal speech input, then deaf adults may still be able to acquire this knowledge. To investigate whether these deaf adults differ from hearing adults in the use of orthographic structure, the performance of a group of normally-hearing subjects was compared with that of deaf subjects.

The use of orthographic structure was investigated testing recognition and recall of fingerspelled letter strings. Fingerspelling is a manual communication system based on English in which words are spelled out by the sequential production of letters of the manual alphabet. As shown in Figure 1, the American manual alphabet has a handshape for each letter of the English alphabet. Fingerspelling is used both in American Sign Language (ASL or Ameslan) and in manual communication systems based on English.

In fingerspelling, words are presented as a temporally sequential display of individually produced letters with an average presentation rate of 20 msec per letter (Bornstein, 1965). Letters are displayed with the hand held in one spatial location. For printed letters, display characteristics such as this make word recognition difficult. With sequential presentation of printed letters displayed in one spatial location, normally-hearing readers can accurately name words only when the duration of each letter is at least 375 msec (Kohlers & Katzman, 1966). Even when the printed letters are spatially distinct, ability to read words is dramatically reduced for sequentially displayed individual letters compared with multi-letter displays (Newman, 1966). Fingerspelling provides an interesting case in word recognition in that fingerspelled words can be recognized at rates that are difficult for the recognition of sequentially presented printed letters. For this reason, a secondary goal of the present research was to examine skilled reading of fingerspelling.

A sequential presentation of letters might suggest sequential recognition of individual letters. However, it may be that, similar to the recognition of printed words, orthographic structure is used in the recognition of fingerspelled words. Since it has been demonstrated that there are "co-articulatory" effects in skilled fingerspelling, with letter context influencing letter production (Reich, 1974), this could allow for the use of orthographic structure in word recognition.

In the present experiment, fingerspelled words, pseudowords, and nonwords were presented to deaf and hearing adults skilled in the use of fingerspelling. If orthographic structure is used in processing the fingerspelled stimuli, then letters of orthographically regular nonsense words should be recalled more accurately than letters of orthographically irregular nonsense words. Errors in letter report for words were analyzed to examine orthographic regularities in production for both deaf and hearing subjects.

METHOD

Stimuli

Sixty stimulus items were used. Thirty were real words chosen from samples of words found misspelled in writing by deaf adults. These words ranged in length from five to thirteen letters, mean length being 8.3 letters per word. The words ranged in frequency of occurrence from 1 - 190 (median of 10.5) according to Kucera and Francis (1967). These thirty words were matched in mean length with 20 orthographically regular pseudowords (e.g., BRANDIGAN, MUNGRATS, VISTARMS) and 10 orthographically irregular nonwords (e.g., FTERNAPS, PKANT, VETMFTERN). The selection criteria for the orthographically regular and irregular words were in accord with the criteria outlined in Appendix A of Massaro, Venezky, and Taylor (1979). According to these criteria, the regular strings (pseudowords) were pronounceable and had ortho-

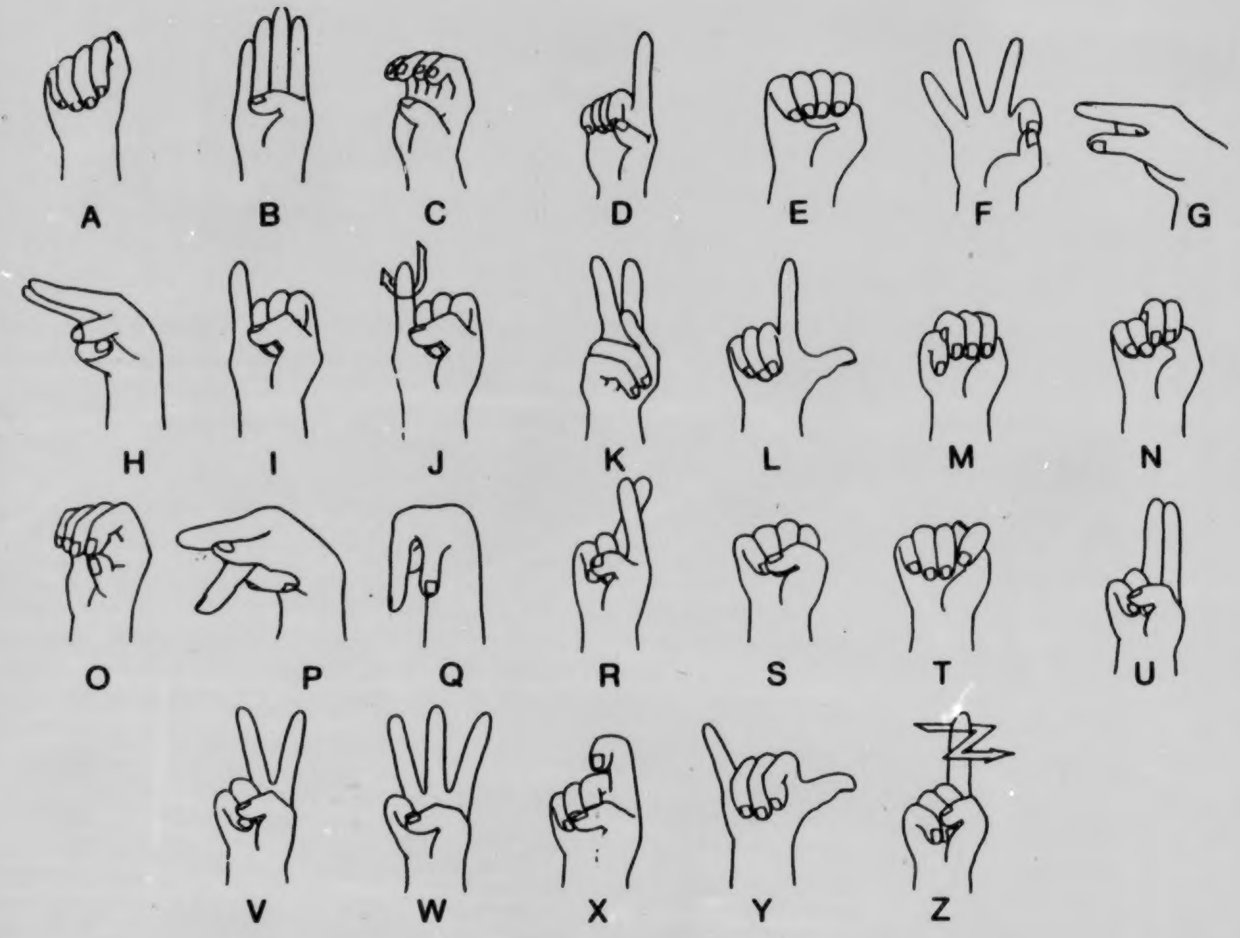

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

Figure 1. Handshapes of the American Manual Alphabet.

graphically legal spelling patterns. The irregular strings (nonwords) contained unpronounceable consonant clusters. A complete listing of the stimuli is given here in the Appendix.

Stimuli were recorded on videotape by a deaf native signer of ASL (i.e., a person who had deaf parents and had learned ASL as a first language). The signer made no mouth movements nor facial expressions that would indicate the lexical status of items. Measurement of the length of each recorded item revealed a mean presentation rate of 354 letters per minute. This rate is consistent with the rate found by Bornstein (1965) to be a natural ASL rate. The production rate for the thirty words did not differ from the production rate for the other thirty items, $\underline{t}(58)=1.87$, $\underline{p}>.05$. Words, pseudowords, and nonwords were mixed throughout the list. Following each item, a blank interval of approximately 10 seconds was recorded for use as a response period.

Procedure

Subjects were instructed that they would see many fingerspelled items and that for each they were to make two responses: First, write the letters of the item they had just seen; second, make a lexical decision. They were to circle YES or NO on their answer sheet to indicate whether they thought the presented letter string was or was not an actual word. The instructions, signed in ASL, were recorded on videotape.

Subjects were run in groups of one to six persons. The entire experiment lasted approximately 30 minutes.

Subjects

A group of deaf subjects and a group of hearing subjects were tested. Subjects in both groups had deaf parents and had learned fingerspelling from their parents.

Deaf subjects were 14 congenitally deaf adults recruited through New York University and California State University, Northridge. All were profoundly deaf. There were six women and eight men, ranging in age from 17 - 53 years, median age 28.5 years.

Hearing subjects were recruited through interpreter services in Connecticut and New York. There were five women and three men ranging in age from 22 - 49 years, median age 29 years.

RESULTS AND DISCUSSION

To examine possible processing differences for the two groups, the eight hearing subjects were matched in overall accuracy with eight deaf subjects. Overall accuracy was determined for each subject as the percentage of correct responses across conditions. Only items for which there was both a correct lexical decision and a correct report of all letters were considered to be correct responses. Overall, hearing subjects had an accuracy rate of 43.7% (range 21.7% - 65.0%). Eight deaf subjects, whose accuracy was in the middle of the performance range for the 14 who participated, performed at a comparable level. Overall they were 40.8% accurate (range 20.0% - 70.0%), which was not significantly different from the accuracy of the hearing subjects, $t(14)=.34$, $p>.05$. Further analyses are based on these two matched groups of eight subjects each.

Lexical Identification

Mean overall accuracy for the lexical decision task was 85.5%. Analysis of group (deaf, hearing) by stimulus type (words, pseudowords, nonwords) found that there was no significant difference in accuracy across stimulus type, $F(2,28)=.88$, $MS_e=70.06$, $p>.05$, nor was there an interaction between group and stimulus type, $F(2,28)=.74$, $MS_e=70.06$, $p>.05$. There was a tendency for deaf subjects to perform this task more accurately than hearing subjects, although the difference only approached significance, $F(1,14)=3.00$, $MS_e=339.35$, $p<.20$. The performance of both groups of subjects in this task is shown in Table 1.

Table 1

Mean percentage correct lexical decisions and correct identification of words.

	Deaf	Hearing
Lexical Decisions		
Words	91.7%	80.4%
Pseudowords	89.4%	84.4%
Nonwords	88.8%	77.5%
Word Identification| correct lexical decision	93.0%	92.6%

To ensure that this high accuracy could not have been due to some non-linguistic cue to wordness of the stimulus items (e.g., facial cues or "awkward" production of pseudowords and nonwords), eight hearing adults, naive with respect to fingerspelling, were asked to make lexical decisions regarding the stimuli. They viewed the videotape and were told that for every fingerspelled item they were to circle YES or NO on their answer sheet to indicate whether or not they thought the item was an actual word. This group of naive adults was only 49.2% accurate in the task, a rate that does not differ from chance performance, $\chi^2(1)=.05$, $p>.05$. Therefore, the high accuracy of the two groups of deaf and hearing subjects in this task can be attributed to their knowledge of fingerspelling.

The whole report technique of the present experiment allowed for a determination as to whether or not there was correct identification of words. Three types of response errors were considered to be failures to identify the word. First were those responses on which more than 50% of the letters were omitted. These omissions were rare; only three such errors were made (by deaf subjects). The second source of error consisted of responding with a morphologically incorrect form of the word (e.g., *baptized* for BAPTIZE) and accounted for three errors of the deaf subjects and five errors of the hearing subjects. The third source of error consisted of responding with the wrong word (e.g., *complicate* for COMMUNICATE), accounting for five errors of the deaf subjects and nine errors of the hearing subjects. Table 1 presents subjects' accuracy of word identification given a correct lexical decision. Deaf and hearing subjects did not differ in their accuracy on word identification, $t(14)=.09$, $p>.05$.

These latter two sources of error in lexical identification appear to result from guessing the word on the basis of a few letters. It should be pointed out that this strategy of identifying a word on the basis of a few perceived letters is not a bad one in normal conversations. In these conversations, letters of fingerspelled words are often omitted or sloppily produced (Caccamise, Hatfield, & Brewer, 1978), but within the syntactic and semantic context provided by the conversation, word identification from partial information is possible. In the present task, however, recognition of only a few letters led to the errors in lexical identification. These errors resulted both in incorrectly identifying actual words and in incorrectly responding that pseudowords and nonwords were words (e.g., *raps* for RAPAS and *veteran* for VETMFTERN).

Letter Report Accuracy

Given a correct response on the lexical decision task and a correct lexical identification of the words, how accurate were subjects at reporting all the letters of an item? A summary of this performance by the two groups on each word type is shown in Table 2. An analysis of the percentage correct was performed on group (deaf or hearing) by word type (word, pseudoword, nonword) for trials on which there was a correct lexical decision and identification. The analysis revealed a strong effect of word type, $F(2,28)=170.03$, $MS_e=129.32$, $p<.001$. This difference was significant between all word types (Newman-Keuls, $p<.01$), thus indicating effects of word familiarity (letters of words better recalled than letters of pseudowords) and orthographic structure (letters of pseudowords better recalled than letters of nonwords). There was no main effect of group for accuracy of letter report, $F(1,14)=1.65$, $MS_e=368.33$, $p>.05$, but there was an interaction of group by word type, $F(2,28)=6.70$, $MS_e=129.32$, $p<.005$. Analysis of the simple effects revealed that the two subject groups differed in letter report accuracy for words, $F(1,28)=13.93$, $p<.001$, but did not differ significantly in letter report accuracy for pseudowords, $F(1,28)=.00$, $p>.05$, or nonwords, $F(1,28)=.29$, $p>.05$. Thus, the interaction of group by word type was due to greater accuracy by hearing than deaf subjects on letter report for words.

Table 2

Mean percentage correct report of all letters given a correct identification of words and a correct lexical decision for the pseudowords and nonwords.

	Deaf	Hearing
Words	70.2%	92.3%
Pseudowords	30.7%	31.4%
Nonwords	9.3%	6.0%

If subjects were using orthographic structure in the processing of words and pseudowords, there should be nonindependence of letter report for these stimuli. That is, the probability of letter report of a given letter should be a function of the probability of the recall of the other letters in the word. This interdependence of letter report would not be expected to be involved in letter report for nonwords, however, since principles of English orthography are violated in these nonwords. Tests for independence of letter report were performed separately on words, pseudowords, and nonwords. Independence is indicated if the following equation holds:

$$p(\text{all letters of an item}) = p(\text{individual letter})^n \quad (1)$$

where n=number of letters in the word.

Analyzing for group (deaf or hearing) by word length by probability (all letters vs. individual letters), it was found that for words and pseudowords the probability of correctly reporting all the letters of the item was greater than the probability of reporting the letters independently: for words, $F(1,14)=71.71$, $MS_e=263.40$, $p<.001$; for pseudowords, $F(1,14)=26.95$, $MS_e=285.02$, $p<.001$. The effect did not interact with subject population for either the analysis of words, $F(1,14)=.22$, $MS_e=263.40$, $p>.05$, or pseudowords, $F(1,14)=.81$, $MS_e=285.02$, $p>.05$. Thus, the letters of words and pseudowords were not processed independently. However, for the fingerspelled items that violated orthographic structure (the nonwords), the probability of correctly reporting all the letters of the item was not greater than the probability of independently reporting each letter, $F(1,14)=1.98$, $MS_e=103.43$, $p>.05$. For nonwords, therefore, the letters were processed independently. As before, there was no interaction with subject group, $F(1,14)=1.35$, $MS_e=103.43$, $p>.05$.

These results give evidence for the ability of deaf adults to use orthographic structure. Similar to the orthographic structure effects previously reported for deaf adults by Gibson et al., (1970), the present study found greater accuracy in letter report for pseudowords than nonwords. In accord with these findings, the nonindependence of letter processing for words and pseudowords indicates interdependence of letter processing. That is, processing of a given letter was influenced by other letters of the word or pseudoword. There were also other indications that deaf and hearing subjects in the present experiment were aware of violations of English orthography: When the fingerspelled nonwords were presented, subjects often laughed. Generally a look of surprise would appear on their faces at these violations of orthographic structure.

Together, the above results also clearly indicate that orthographic structure is used in processing fingerspelling. They indicate that even though letter presentation is temporally sequential, letter processing is influenced by surrounding letters. Since there are coarticulatory effects in skilled fingerspelling, it is reasonable to assume that a fingerspelled letter contains information about adjacent letters. A skilled fingerspeller would, therefore, be expected to make use of this context information in word recognition (see Wickelgren [1969, 1976] for a discussion of context-sensitive coding in speech). This context-sensitivity may explain how orthographic structure is able to be used in identifying fingerspelled letters despite the temporally sequential display of letters, and may explain how these sequential letters can be processed so much more rapidly than sequentially presented printed letters.

In addition to the effects of orthographic structure, word familiarity effects were found here. These word familiarity effects, involving better recall of letters of words than letters of pseudowords, are consistent with the greater accuracy of letter report for letters of printed words than for letters of printed pseudowords (Manelis, 1974; Spoehr & Smith, 1975). A word superiority effect of fingerspelled words over fingerspelled nonwords, consistent with the present findings, has been reported earlier by Zakia and Haber (1971).

Some theorists attribute the word familiarity effect to the fact that words allow for holistic recognition of visual configurations. However, it is unlikely that this interpretation can account for the present results for the following reason: The majority of stimulus words would rarely, if ever, have been seen as fingerspelled words by the subjects prior to this experiment because the words would tend to be signed rather than fingerspelled in signed conversations. The familiarity that the subjects have with these words, therefore, would be with the printed form of the word. This situation is analogous, perhaps, to that of the recognition of mixed-case printed words in that the orthographic integrity of the words is preserved, but the visual configuration is disrupted. Studies have found the while there is a perceptual advantage for same-case over mixed-case words, word familiarity effects are obtained with mixed-cases, indicating that the word familiarity effect need not be totally attributable to holistic word recognition (Coltheart & Freeman, 1974; McClelland, 1976). What, then, contributes to superior letter report for words in the present experiment?

Two factors appear to be involved. The first is that the associations between letter sequences of words should be stronger than the associations for the sequences of permissible although novel items. These stronger associations would allow more perceptual facilitation of letters for words than pseudowords (Adams, 1979). Thus, the letters would be more accurately recognized for words than for pseudowords. The second factor contributing to the word familiarity effect is one of memorability. Pseudowords and nonwords

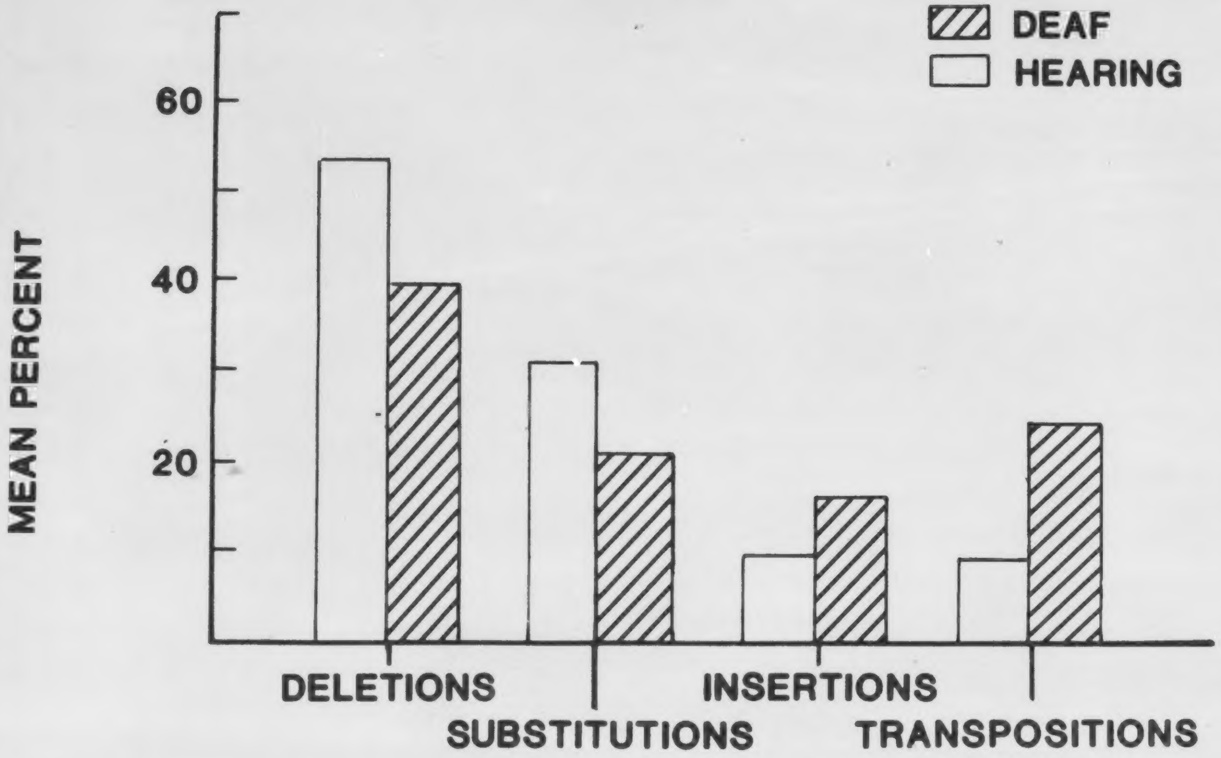


Figure 2. Mean percentage occurrence of each error type for deaf and hearing subjects.

represent novel letter sequences. The subjects must recall the letters based on a single presentation. But for words, once the word is recognized, the subjects are able to bring their productive spelling abilities to bear on the task of letter report. Incorrect letter reports for these words, in this respect, represent spelling errors.

Incorrect Letter Reports for Words

Each incorrect letter report for a correctly identified word was scored in three ways: (1) each was classified as to whether or not the reported letter string produced a sequence that preserved the phonetic structure of the presented word, (2) each was classified as to whether the reported sequence was orthographically regular or irregular, and (3) each was classified as to the type of error.

For hearing adults and children, the predominant form of spelling error is a phonetically consistent but orthographically incorrect rendering of the intended word (Fischer, 1980; Frith, 1980; Masters, 1927). In these misspellings, each phonetic segment of the word is graphemically represented in the order of occurrence. The phonetic structure is therefore maintained in the misspelling. Did the incorrect responses for words in the present experiment preserve the phonetic structure of the words presented? Analysis revealed that the hearing subjects made more incorrect letter reports phonetically equivalent to the target word than did the deaf subjects, $\chi^2(1) = 10.01$, $\underline{p} < .005$. For hearing subjects, the mean percentage of such responses was 63.6%; for deaf subjects, the corresponding percentage was only 18.6%.

But while the letter reports for deaf subjects were not consistent with the phonetic structure of the target word, by and large, the responses were orthographically regular. Orthographically regular words, in accord with Massaro et al. (1979), were both pronounceable and contained only legal consonant and vowel clusters. For deaf subjects, 93.8% of the incorrect responses were regular English letter sequences. For hearing subjects, 95.8% were regular. There was no difference in the frequency of deaf and hearing subjects making such responses, $\chi^2(1) = .62$, $\underline{p} > .05$. This indicates that, like hearing adults, deaf adults have a definite knowledge of English orthographic constraints.

The incorrect letter reports were further classified using the following categories of error type: letter deletions, substitutions, insertions, and transpositions. As shown in Figure 2, a major difference in error type for deaf and hearing subjects was the tendency for deaf, but not hearing, subjects to order the letters of words incorrectly, resulting in a transposition of phonetic segments. Some examples are adveristement for ADVERTISEMENT, funreal for FUNERAL, hemsiphere for HEMISPHERE, viedo for VIDEO, and vechile for VEHICLE. While deaf subjects made 17 errors of this type (representing 23.9% of their total errors), only 1 such error was made by hearing subjects (7.7% of the total). For hearing persons, the incidence of letter transpositions is generally so low that it may be possible to account for all the errors in spelling experiments without even including a category for letter transpositions (Fischer, 1980). It is interesting to note that none of these transpositions preserved the phonetic structure of the words. In all cases the transposed letters incorrectly ordered phonetic segments.

Little work has been undertaken to understand the spelling process for deaf persons. The present finding of the low percentage of letter reports phonetically equivalent to the target is of great interest as it suggests that the cognitive processes underlying spelling for deaf adults may be fundamentally different from those underlying productive spelling for hearing adults. Models of the spelling process for hearing persons commonly hypothesize that productive spelling involves generating a phonetic representation of the target word and then generating possible orthographic realizations of this representation (Frith, 1980; Simon & Simon, 1973). These models therefore account for the tendency of hearing persons to make misspellings that preserve the phonetic structure of the intended word.

The low incidence of phonetically consistent letter reports by the deaf adults in this experiment suggests that the spelling process of deaf persons is not well described by these models. The few studies that have been concerned with the spelling process for deaf persons have been conducted with deaf children. Dodd (1980) examined the spelling of words by orally-trained deaf children in England. The task in Dodd's experiment was to lipread words pronounced by the experimenter and then spell the words. The children (mean age 14.5 years) made only about 11% misspellings that were classified as reflecting the phonetic structure of the pronounced words. It should be noted, however, that 64.8% of the children's errors were classified as "refusals" to spell the pronounced word. If only the actual misspellings of the children are considered in Dodd's data, the incidence of phonetically consistent misspellings is 31.5%.

In another experiment designed to determine the underlying spelling processes of deaf children, Hoemann, Andrews, Florian, Hoemann, and Jansema (1976) found that few of the misspellings of the deaf and hearing-impaired children they tested could be considered phonetically equivalent to the target word. Children in their experiment were ages 6 - 19 years and were being educated with the Rochester Method (combined speech and fingerspelling). The children were presented with pictures of objects and were asked to spell the name of each of the objects. Earlier work had found that the majority of misspellings made by hearing children on this task were phonetically consistent with the target word (Mendenhall, 1930). No more than 19% of the misspellings of the children tested by Hoemann et al. could be considered phonetically equivalent to the target.

Cromer (1980) analyzed samples of free writing from six orally-educated deaf children in England (median age 10.5 years). By Cromer's analysis, 62.25% of the misspellings of the deaf children were "phono-graphic errors," defined as resembling "in some respect the sound of the target word when pronounced" (p. 412). By this definition, errors such as basking for "basket" and amanals for "animals" were scored as phono-graphic errors. Thus, not all these phono-graphic errors would be phonetically consistent with the target. Examining Cromer's corpus of errors according to the present concern of phonetic misspellings, it is apparent that only few of the misspellings can be considered to be consistent with the phonetic structure of the intended word.

One final point is worth mentioning. The deaf children studied by Cromer (1980) made transpositions similar to those made by the adults in the present study. Cromer classified these transpositions and ordering errors under the category "Visual Errors." While this type of error accounted for 15.75% of the errors made by the deaf children in his study, no such errors were made by the normally-hearing control children. As with the present findings, these ordering errors did not preserve the phonetic structure of the target word.

Together, these studies with children and the present one with adults are consistent, a consistency that is especially striking given the differences in methodology of these studies and the differences in the age and language background of the subjects. These studies converge on the finding that deaf persons do not make phonetically consistent spelling errors to the degree that hearing persons do. The suggestion from these findings is that the spelling process for deaf persons may be fundamentally different from the spelling process for hearing persons. In comparison with hearing persons, it appears that deaf persons may make less use of a stored phonetic representation of words when spelling. The types of errors they make appear to be consistent with what Ellis (in press) terms as errors based on "partial lexical knowledge" in which the speller knows "some but not all of the letters, or all of the letters but not the correct order." While the errors of deaf adults in the present experiment were often consistent with this definition, it should be borne in mind that deaf signers may have additional strategies available to them.

One strategy of deaf signers in the present experiment deserves mention: Very often the deaf subjects fingerspelled items on their hands before writing their responses. This fingerspelling often allowed them to try different spellings in an effort to decide the correct letter sequence. This "trying out" spellings should not be thought of as equivalent to the strategy of writing down various spellings of a word to determine which "looks" correct. The deaf subjects occasionally employed this strategy also. Rather, the manual strategy seems more to determine which spelling "feels" correct. Often subjects did not even look at their hands while trying out the letter sequences. Often the hand they were not using for writing their answers was held under the table as they fingerspelled the different sequences. This suggests, therefore, that whereas a component of the spelling process for hearing persons is phonetic, a component of the spelling process for deaf signers may be kinesthetic.

CONCLUSIONS

The present experiment clearly indicates that deaf adults are able to make use of orthographic structure. This was shown both in the recall advantage for orthographically regular over orthographically irregular letter strings, similar to the findings of Gibson et al. (1970), and in the analysis of errors in letter reports for words. These results support the conclusion of Gibson et al. that while the mapping of speech sounds to graphemes may facilitate the acquisition of orthographic structure for hearing persons, congenitally and profoundly deaf persons are, nevertheless, able to acquire this knowledge.

The analysis of letter report errors for words indicates that the deaf adults were sensitive to the orthographic regularities of English in their productions. But in contrast to the hearing adults, the responses of deaf adults were not consistent with the phonetic structure of words. These results suggest that deaf adults may use orthographic but not phonetic structure when spelling.

A secondary goal of the present research was to examine recognition of fingerspelled words. The work suggests that signers, both deaf and hearing, make use of orthographic structure in the processing of fingerspelling.

REFERENCES

Adams, M. J. Models of word recognition. Cognitive Psychology, 1979, 11, 133-170.

Aderman, D., & Smith, E. E. Expectancy as a determinant of functional units in perceptual recognition. Cognitive Psychology, 1971, 2, 117-129.

Baron, J., & Thurston, I. An analysis of the word-superiority effect. Cognitive Psychology, 1973, 4, 207-228.

Bornstein, H. Reading the manual alphabet. Washington, D.C.: Gallaudet College Press, 1965.

Caccamise, F., Hatfield, N., & Brewer, N. Manual/Simultaneous Communication (M/SC) research: Results and implications. American Annals of the Deaf, 1978, 123, 803-823.

Carr, T. H., Davidson, B. J., & Hawkins, H. L. Perceptual flexibility in word recognition: Strategies affect orthographic computation but not lexical access. Journal of Experimental Psychology: Human Perception and Performance, 1978, 74, 674-690.

Chomsky, N. Phonology and reading. In H. Levin & J. P. Williams (Eds.), Basic studies on reading. New York: Basic Books, 1970.

Coltheart, M., & Freeman, R. Case alternation impairs word identification. Bulletin of the Psychonomic Society, 1974, 3, 102-104.

Cromer, R. F. Spontaneous spelling by language-disordered children. In U. Frith (Ed.), Cognitive processes in spelling. London: Academic Press, 1980.

Dodd, B. The spelling abilitites of profoundly pre-lingually deaf children. In U. Frith (Ed.) Cognitive processes in spelling. London: Academic Press, 1980.

Doehring, D. G., & Rosenstein, J. Visual word recognition by deaf and hearing children. Journal of Speech and Hearing Research, 1960, 3, 320-326.

Ellis, A. W. Spelling and writing (and reading and speaking). In A. W. Ellis (Ed.), Normality and pathology in cognitive functions. London: Adademic Press, in press.

Fischer, F. W. Spelling proficiency and sensitivity to linguistic structure. Unpublished doctoral dissertation, University of Connecticut, 1980.

Frith, U. Unexpected spelling problems. In U. Frith (Ed.), Cognitive processes in spelling. London: Academic Press, 1980.

Gibson, E. J., Pick, A., Osser, H., & Hammond, M. The role of grapheme-phoneme correspondence in the perception of words. American Journal of Psychology, 1962, 75, 554-570.

Gibson, E. J., Shurcliff, A., & Yonas, A. Utilization of spelling patterns by deaf and hearing subjects. In H. Levin & J. P. Williams (Eds.), Basic studies in reading. New York: Basic Books, Inc., 1970.

Hoemann, H. W., Andrews, C. E., Florian, V. A., Hoemann, S. A., & Jansema, C. J. The spelling proficiency of deaf children. *American Annals of the Deaf*, 1976, *121*, 489-493.

Klima, E. S. How alphabets might reflect language. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye*. Cambridge, Mass.: MIT Press, 1972.

Kohlers, P. A., & Katzman, M. T. Naming sequentially presented letters and words. *Language and Speech* 1966, *9*, 84-95.

Kucera, H., & Francis, W. *Computational analysis of present-day American English*. Providence, R.I.: Brown University Press, 1967.

Manelis, L. The effect of meaningfulness in tachistoscopic word perception. *Perception & Psychophysics*, 1974, *16*, 182-192.

Massaro, D. W., Venezky, R. L., & Taylor, G. A. Orthographic regularity, positional frequency, and visual processing of letter strings. *Journal of Experimental Psychology: General*, 1979, *108*, 107-124.

Masters, H. V. *A study of spelling errors*. Iowa City: University of Iowa Studies in Education, IV (4), 1st Series 138, September 1, 1927.

McClelland, J. L. Preliminary letter identification in the perception of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, *1*, 80-91.

Mendenhall, J. E. *An analysis of spelling errors: A study of factors associated with word difficulty*. New York: Bureau of Publication, Teachers College, Columbia University, 1930.

Newman, E. B. Speed of reading when the span of letters is restricted. *American Journal of Psychology*, 1966, *79*, 272-278.

Reich, P. A. Visible distinctive features. In A. Makkai & V. B. Makkai (Eds.), *The First LACUS Forum*. Columbia, S.C.: Hornbeam Press, 1974, 348-356.

Simon, D. P., & Simon H. A. Alternative uses of phonemic information in spelling. *Review of Educational Research*, 1973, *43*, 115-137.

Spoehr, K., & Smith, E. The role of orthographic and phonotactic rules in perceiving letter patterns. *Journal of Experimental Psychology: Human Perception and Performance*, 1975, *1*, 21-34.

Tenney, Y. J. Visual factors in spelling. In U. Frith (Ed.), *Cognitive processes in spelling*. London: Academic Press, 1980.

Venezky, R. *The structure of English orthography*. The Hague: Mouton, 1970.

Wickelgren, W. A. Context-sensitive coding, associative memory, and serial order in (speech) behavior. *Psychological Review*, 1969, *76*, 1-15.

Wickelgren, W. A. Phonetic coding and serial order. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception, Vol VII: Language and Speech*. New York: Academic Press, 1976.

Zakia, R. D., & Haber, R. N. Sequential letter and word recognition in deaf and hearing subjects. *Perception & Psychophysics*, 1971, *9*, 110-114.

Appendix

Words

ADVERTISEMENT
AWKWARDLY
BANKRUPTCY
BAPTIZE
CADILLAC
CAREFUL
CHIMNEY
COMMUNICATE
ELABORATE
FUNERAL
GRADUATE
HELICOPTER
HEMISPHERE
INTERRUPT
MOUNTAIN
PANTOMIME
PHILADELPHIA
PHYSICS
PREGNANT
PSYCHOLOGICAL
PUMPKIN
RHYTHM
SUBMARINE
SURGERY
THIRD
TOMATO
UMBRELLA
VEHICLE
VIDEO
VINEGAR

Pseudowords

BRANDIGAN
CADERMELTON
CHIGGETH
COSMERTRAN
EAGLUMATE
FREZNIK
FRUMHENSER
HANNERBAD
INVENCHIP
MUNGRATS
PHALTERNOPE
PILTERN
PINCKMORE
PRECKUM
RAPAS
SNERGLIN
STILCHUNING
SWITZEL
VALETOR
VISTARMS

Nonwords

CONKZMER
ENGKSTERN
FTERNAPS
HSPERACH
PGANTERLH
PIGTLANING
PKANT
RANGKPES
RICGH
VETMFTERN

EXPLORING THE INFORMATION SUPPORT FOR SPEECH*

J. A. Scott Kelso+ and Betty Tuller++

Abstract. A well-established feature of speech production is that talkers, faced with both anticipated and unanticipated perturbations, can spontaneously adjust the movement patterns of articulators such that the acoustic output remains relatively undistorted. Less clear is the nature of the underlying process(es) involved. In this study we examined five subjects' production of the point vowels /i, a, u/ in isolation and the same vowels embedded in a dynamic speech context under normal conditions and a combined condition in which (a) the mandible was fixed by means of a bite block, (b) proprioceptive information was reduced through bilateral anaesthetization of the temporomandibular joint, (c) tactile information from oral mucosa was reduced by extensive application of topical anaesthetic, and (d) auditory information was masked by white noise. Analysis of formant patterns revealed minimal distortion of the speech signal under the combined condition. These findings are unfavorable for central (e.g., predictive simulation) or peripheral closed-loop models, both of which require reliable peripheral information; they are more in line with recent work suggesting that movement goals may be achieved by muscle collectives that behave in a way that is qualitatively similar to a nonlinear vibratory system.

The remarkable generativity of human movement is a mystery that continues to resist explanation. Within limits, people (and animals) can achieve the same 'goal' through a variety of kinematic trajectories, with different muscle groups and in the face of ever-changing postural and biomechanical requirements. This phenomenon--variously referred to as motor equivalence (Hebb, 1949) or equifinality (von Bertalanffy, 1973)--has been demonstrated again quite recently by Raibert (1978), who showed writing patterns to be characteristic of the same individual even when produced by structures (such as the foot or mouth) that had never previously been used for the act of writing.

---

*A preliminary version of this paper was presented at the 101st meeting of the Acoustical Society of America, May 18-22, 1981.

+Also Departments of Psychology and Biobehavioral Sciences, The University of Connecticut.

++Also Department of Neurology, New York University Medical Center.

Acknowledgment. This research was supported by NIH Grants NS13617 and Biomedical Research Support Grant RR05596 to Haskins Laboratories and NIH Post-Doctoral Fellowship 1-F32-NS-06718-01 to the second author. We thank Thomas Gay and the University of Connecticut Health Center at Farmington for use of facilities. We are especially grateful to Dr. Robert Gross of Louisiana State University Medical Center for performing anaesthetization procedures.

[HASKINS LABORATORIES: Status Report on Speech Research SR-69 (1982)]

Human language is generative in a qualitatively similar way: We seem to have a potentially infinite number of ways of constructing sentences. Nor is it trivial that language, even when stripped of its symbolic component, is a creative or generative activity. Articulatory maneuvers for producing speech sounds can be effected in spite of continuously varying initial conditions. Often the same phonetic segment in different environments can be achieved by very different movement trajectories and end-states. One commonly used experimental paradigm for examining equifinality in speech takes the form of placing a bite block between the teeth, thus fixing the position of the mandible. Under such conditions, so-called "steady state" vowels can be produced apparently without the need for on-line acoustic feedback. Normal range formant patterns are obtained even at the first glottal pitch pulse (Gay, Lindblom, & Lubker, 1981; Lindblom, Lubker, & Gay, 1979; Lindblom & Sundberg, 1971). Moreover, speakers are capable of such "compensatory articulation" with little (if any) articulatory experimentation. Recent work on bite-block speech has shown that response times to produce vowels of the same acoustic quality under normal and bite-block conditions are nearly identical. In addition, the degree of "compensation" (as indexed by deviations from normal formant frequencies) remained unchanged as a function of practice (Fowler & Turvey, 1980; Lubker, 1979). The evidence, then, favors an interpretation that articulatory adjustments to novel contextual conditions are relatively immediate.

What kind of control process could account for the adaptive, generative nature of speech production? An open-loop control system in which commands for producing a given vowel prescribe in detail the activities of relevant muscles can be dispensed with because, by definition, such systems are insensitive to changing contextual conditions. On the other hand, closed-loop control does offer the advantage of adjustment to initial conditions. In peripheral closed-loop, feedback systems, a sensory goal in the form of a spatial (MacNeilage, 1970) or auditory target (Ladefoged, DeClerk, Lindau, & Papçun, 1972; MacNeilage, 1980) is paired with an appropriate set of commands for accomplishing the goal. Resulting sensory consequences are then compared with the sensory goal so that corrections can be made. A potential problem with peripheral closed-loop control is that the corrective process requires time (at least one cycle around the corrective loop). However, if the adjustment to novel conditions is indeed immediate--thus excluding the need for trial and error methods--then a closed-loop mechanism tied to the peripheral motor system fails to capture the phenomenon of interest.

An alternative account favored by Lindblom and colleagues (e.g., Lindblom et al., 1979) replaces the peripheral feedback loop by a central simulation process that derives the expected sensory consequences from a simulated set of motor commands before the actual efferent signals are sent to the periphery. An internal comparison between the simulated and 'target' sensory consequences yields an error signal on the basis of which new (and correct) commands can be emitted. In this manner, adjustments to changes in context can be made in the internal simulation without incurring erroneous effects at the periphery.

It is important to note that the models discussed thus far make the explicit assumption that reliable peripheral information about the articulators' initial conditions is available before motor commands (simulated or actual) are generated. In the peripheral closed-loop model, for example, sensory input must be compared to the internal referent before the output of command signals. In the internal loop model, simulated motor commands are generated for the initial conditions that currently exist (Lindblom et al. 1979). It is not clear in the latter formulation what would happen if contextual conditions changed between the time that simulated and actual motor commands were generated. A more efficient system would be continuously sensitive to, and be capable of modulation by, contextual conditions. For the sake of argument, however, let us assume with Lindblom et al. that one benefit of the internal loop is its speed of correction; possibly the loop is so fast that appropriate output can be generated before contextual conditions have changed.

In any case, for both closed-loop models, elimination or reduction of peripheral information about initial conditions should drastically affect the system's ability to adjust to the novel situation created by a bite block. There are very limited data on this point. Gay and Turvey (1979) found that a single subject (a phonetician) made several attempts before producing 'normal' formant frequencies for the vowel /i/ under conditions in which a bite block was combined with topical anaesthesia of the oral mucosa and bilateral nerve blockage of the temporomandibular joint. Although this result has suggested to some (cf. Perkell, 1979) that joint and tactile information is used to establish an "orosensory frame of reference," we believe there are grounds for caution. One problem is that it is unclear how--given the considerable reduction of peripheral information--Gay and Turvey's subject was capable of adaptive adjustment at all. One possibility, which we consider here, is that auditory information may have played a potentiating role. Although acoustic information does not appear to be a necessary condition for compensatory articulation (e.g., Lindblom et al., 1979), the Gay-Turvey experiment does not preclude an auditory contribution in "recalibrating" the speech system when information from motor structures is rendered unreliable.

The present experiment was designed to examine the role of peripheral information (auditory and somesthetic) in accounts of "immediate adjustment" by asking naive subjects to produce vowels under normal conditions and under bite-block conditions in which somatosensory information was drastically reduced (if not eliminated) and audition was masked by white noise. In addition we address the question of whether the so-called "steady-state" paradigm for bite-block vowels reflects normal dynamic speech motor processes. By examining the production of vowels embedded in a dynamic speech context as well as in isolation, we can discover what differences there are, if any, in observed acoustic patterns. As we shall see, the availability of peripheral information from neither auditory nor peripheral motor structures appears to be crucial to immediate adjustment. We take this result as non-supportive for extant models of the phenomenon. In their place, we offer a class of model--emerging in other areas of motor control (Bizzi, 1980; Fel'dman, 1966, 1980; Kelso, 1977; Kelso & Holt, 1980; Kelso, Holt, Kugler, & Turvey, 1980; Polit & Bizzi, 1978) as well as in the recent speech production literature (cf. Fowler, 1977; Fowler, Rubin, Remez, & Turvey, 1980)--that identifies functional groupings of muscles as exhibiting properties qualitatively similar to a nonlinear oscillatory system. The bottom line of this model and of the present paper is that the equifinality characteristic of vowel production may not be prescribed by closed-loop servomechanisms of the peripheral or central kind. Rather, we argue that it may be a consequence of the parameterization of a dynamical system whose design is intrinsically self-equilibrating. That is, a design in which equilibrium points are a natural by-product of the stiffness and damping specifications for the vowel-producing system.

Subjects. Four female volunteers were paid to participate in this experiment. All were naive to the purpose of the experiment. A fifth subject (male) who was phonetically trained and had prior experience in a similar experiment (see Gay & Turvey, 1979) was also included.

Stimuli. The subjects' task was to say the point vowels /i, a, u/ in isolation and in a /p/-vowel-/p/ context. The /pVp/ syllables were spoken in the carrier phrase "A ___ again." Utterances were produced in three groups of three tokens of a particular vowel or phrase. The subjects were instructed to produce all tokens of a given utterance in exactly the same fashion, with a clear pause after each token. They were also told not to talk between experimental conditions or to practice the production task.

Conditions. The bite block used was a small acrylic cylinder with wedges carved out of each end so that it could fit snugly between the teeth. A 5 mm bite block was used to restrict the normally low jaw position for production of /a/ and /pap/. Either a 17 mm or a 23 mm bite block was used (depending on the individual subject's oral dimensions) for production of /i, u, pip, pup/, which normally involve a high jaw position.

All anaesthetic procedures were performed by Dr. Robert Gross, a specialist in oral and maxillofacial surgery who had collaborated with us in earlier work (Tuller, Harris, & Gross, 1981). Tactile information from the oral mucosa was reduced by spraying the surface of the tongue and oral cavity with a 2% Xylocaine solution. The effectiveness of the topical anaesthesia was tested by pricking the surfaces with a needle until the subject no longer reported sensation. A few catch trials were also included in an attempt to insure honest reporting on the part of the subject. Information from mechanoreceptors in the jaw was reduced by injecting percutaneously a 2% Xylocaine solution directly into left and right temporomandibular joint capsules to achieve auriculotemporal nerve blockage. Chemical blockage of this nerve drastically impairs perception of joint position and movement (cf. Thilander, 1961). This condition will be referred to as the TMJ block.

In order to restrict the availability of auditory information, white noise was presented to the subject over headphones at approximately 90 dB. The subject was told to monitor the amplitude of her or his productions by watching a VU meter and to restrict the excursion of the needle to approximately 55 dB or under.

All subjects spoke with and without the bite block prior to the application of anaesthesia and under all experimental conditions. Two of the four naive subjects received the TMJ block before the topical anaesthesia, and the other two subjects underwent topical anaesthesia first. In each of these pairs, one subject spoke under conditions of auditory masking and the other subject was allowed normal auditory information. The phonetically trained subject received topical anaesthesia before the TMJ block and spoke with masking noise in combination with these two procedures.

Measurement procedure. Individual utterances were input through a Ubiquitous spectrum analyzer to a Honeywell DDP-224 computer, using a 12.8 msec window and 40 Hz frequency resolution. The first and second formants of each utterance were measured from a spectral section display. As in previous experiments (e.g., Lindblom et al., 1979; Fowler & Turvey, 1980), acoustic measures of the isolated vowels were made at the first glottal pulse. For many English speakers the isolated vowels may not be truly static, that is, they may show some articulatory movement and thus some shifting of formant frequencies; nevertheless, the adopted procedure was to measure formant frequencies at the first glottal pulse. For the /p/-vowel-/p/ syllables, F1 and F2 values were taken from the point within the vowel at which F2 was most extreme. This point was chosen as the closest approximation to the "target" vowel formants.

Results. The main interest of the present experiment rests on a comparison of speech under normal conditions and conditions of reduced peripheral information. Figure 1 shows the mean values for F1 and F2 for each subject. The top half shows the mean formant values for the isolated vowels, and the bottom half the mean formant values for the /p/-vowel-/p/ syllables. The conditions of speaking are coded as follows: "M" means the subject produced the utterances under conditions of masking noise, "J" is the TMJ block, "T" corresponds to topical anaesthesia, and "BB" is the bite block. Each subject's nine normal productions of a given utterance were compared using t-tests with his or her productions under the most extreme condition of sensory deprivation. None of the subjects (except subject 1) showed any differences in formant frequency values between normal and deprived conditions. Such was the case regardless of whether vowels were spoken in isolation or in a consonantal frame; t(8) values ranged from .05 to 1.79, *p*s >.1. For Subject 1, a significant mean difference occurred between the normal and most deprived condition only for F1 and F2 of the vowel /u/. We are hard put to account for these anomalies: The effect on /u/, though substantial (a 97 Hz difference for F1 and a difference of 363 Hz for F2) is in the direction opposite to expectation. Specifically, the presence of a bite block might be expected to raise all formant frequencies when producing /u/ because of possible structural limitations on lip protrusion and constriction. In contrast, however, this subject's productions of /u/ with a bite block actually showed lower F1 and F2 frequencies than when there was no bite block. Neither can the effect be attributed only to masking (which might implicate higher formant frequencies). Notice that the formant values for combined sensory deprivation conditions are very similar with and without masking. It is also worth remarking that S5 in Figure 1 is the phonetically trained subject whose results conform to the general pattern shown by naive subjects.

Before concluding that these results reflect immediate adjustment in the deprived condition, it is necessary to exclude the possibility that systematic changes in formant values occurred over trials. Figure 2 shows the F1 and F2 values for individual tokens, in order, for the vowel /i/ produced by one subject under the most extreme conditions (i.e. topical anaesthesia, TMJ block, a 23 mm bite block and masking noise). Also shown are the mean F1 and F2 values for this combined condition, and the mean value of the subjects' "normal" formants. The slope of the line formed by tokens one through nine does not differ significantly from the line formed in the (non-bite block) control condition. Evidently, there does not appear to be a systematic

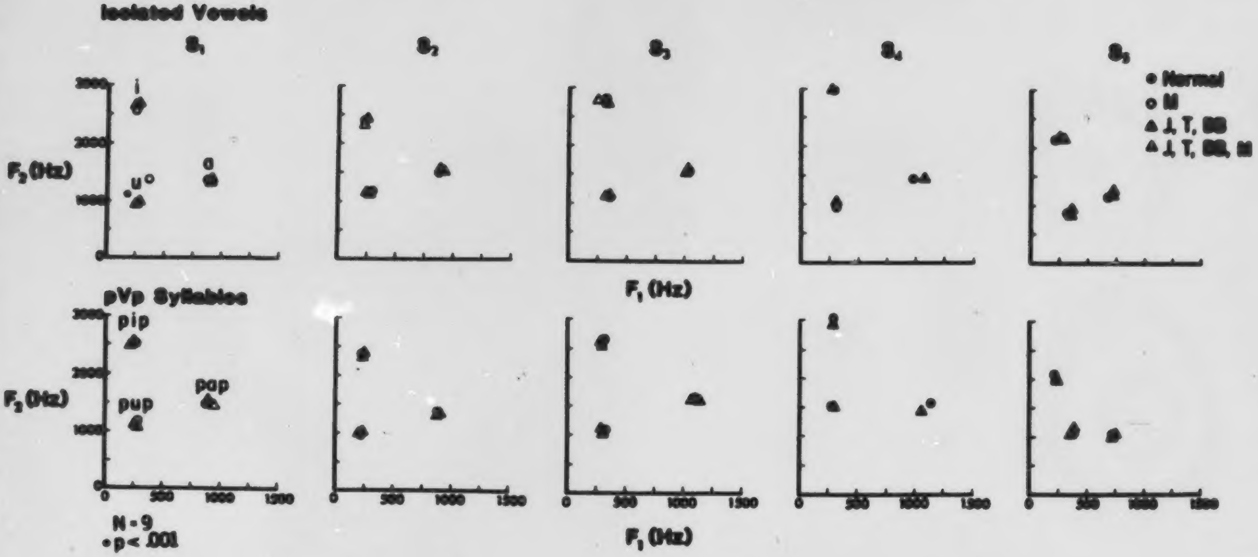


Figure 1. Mean values of F_1 (x-axis) and F_2 (y-axis) for nine repetitions of the indicated utterances. Data for the five subjects are presented separately. Top: Isolated vowels. Bottom: /pVp/ syllables. M = masking noise, J = TMJ block, T = topical anaesthesia, BB = bite block.

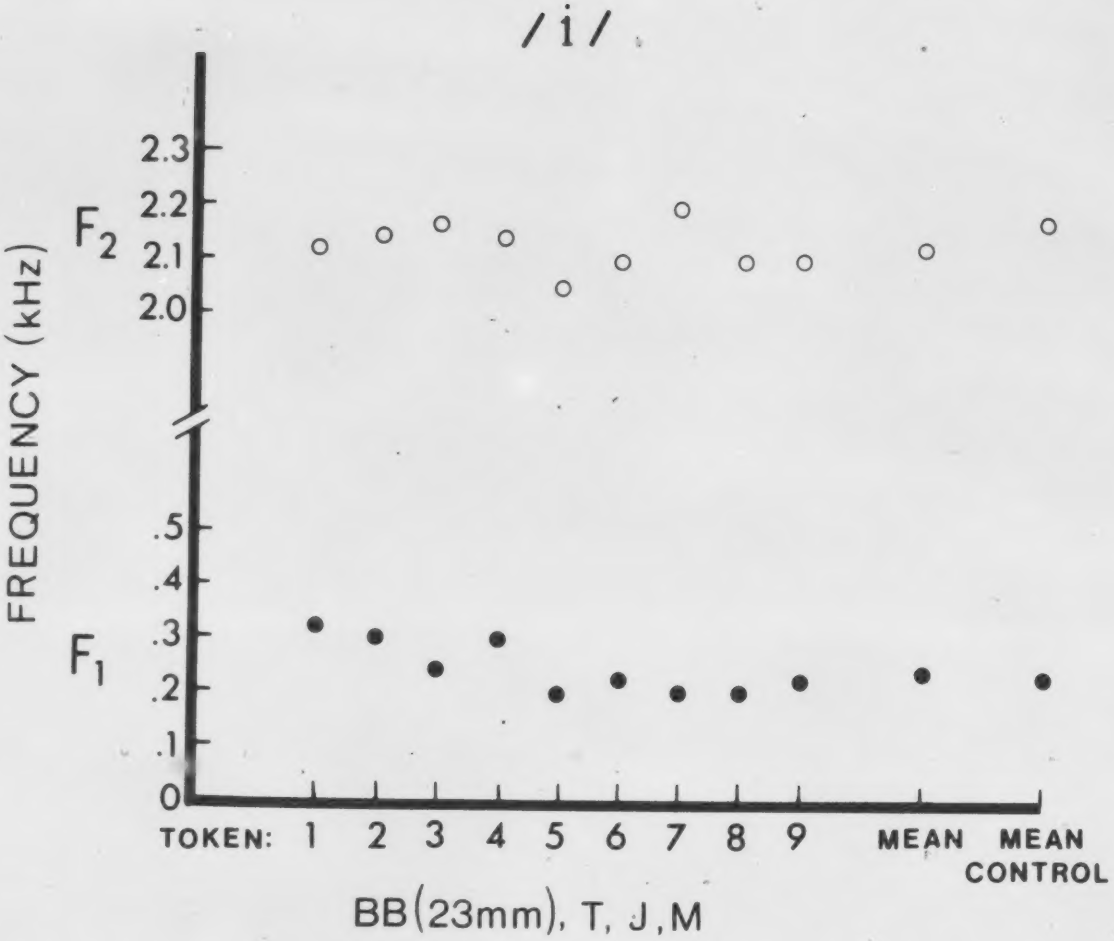


Figure 2. F_1 and F_2 values for individual tokens, in order, of the vowel /i/ produced by one subject under the most extreme experimental condition.

learning effect occurring over trials. We confirmed this statistically for all subjects by performing linear regression analyses across trials of each subject's productions under normal and deprived conditions. Correlations were converted to z-scores and t-tests performed to determine whether the slopes differed between the two conditions. Of the sixty analyses performed (5 subjects by 6 utterance types by 2 formants), not a single one showed a difference in slope; t(7) values ranged from .00 to .84, *p*s >.1.

DISCUSSION

The present data are not easily explained by current models of movement control that have been proposed to account for the remarkable context sensitivity of the speech production system. Closed-loop models--of the central or peripheral kind--both entail an availability of reliable sensory information about the initial conditions of the articulators. However, our experiment shows that acoustically normal vowels can be produced not only when the normal relationships among the articulators are changed by a bite block, but also when sensory information from auditory, joint, and tactile sources is drastically reduced as well. Furthermore, and as other recent work also suggests (cf. Fowler and Turvey, 1980), "articulatory compensation" appears to be achieved immediately and with little or no practice; none of our naive subjects' data provided any evidence of short-term adaptation. In support of the latter claim, the data displayed in Figure 2 are actually from the same subject that apppeared to display motor learning effects in an earlier study (Gay & Turvey, 1979).[1]

Before offering an alternative interpretation of our data in terms other than closed-loop models, two caveats may be in order. The first is that our results do not necessarily refute closed-loop control when the system is in its normal mode, that is, when all sources of information are available. The second is that our paradigm in all likelihood does not completely eliminate peripheral information, and hence a closed-loop simulation model cannot be ruled out completely. Nevertheless, given the drastic reduction in propriospecific information we (and surely the proponents of closed-loop models) might have expected much more severe distortion of the acoustics than was observed here.

In spite of these caveats, we believe that a more parsimonious account of the phenomenon can be forwarded, though it is less well known in speech research than the servoengineering model. The account that we shall consider does not, in fact, depend on whether sensory input about the initial positions of articulators is available or not. Thus there is no requirement for one model when sensory afference is available and another quite different model when it is absent.

The view that we shall express for the present data has been laid out in some detail elsewhere (Fowler et al., 1980; Kelso, Holt, Kugler, & Turvey, 1980). In brief, it argues that functional groupings of muscles--sometimes called synergies (cf. Gelfand, Gurfinkel, Tsetlin, & Shik, 1971) or coordinative structures (cf. Turvey, Shaw, & Mace, 1978)--exhibit behavior qualitatively similar to a (nonlinear) mass-spring system. Such systems are intrinsically self-equilibrating in the sense that the "end-point" of the system, or its "target," is achieved regardless of initial conditions. Thus in normal and deafferented animals (Bizzi, Dev, Morasso, & Polit, 1978), it can be shown that desired limb positions are attainable without starting position information, and even when the limb is perturbed on its path to the target. Similarly, the localization ability of functionally deafferented humans (Kelso, 1977; Kelso & Holt, 1980) and individuals with the metacarpophalangeal joint capsule surgically removed (Kelso, Holt, & Flatt, 1980) is not affected by altered initial conditions or unexpected perturbations. These data have led to the view that the "target" of the system is not achieved by means of conventional closed-loop control; rather it is a consequence of the system's dynamic parameters (mass, stiffness, damping). In such a model, the only parameters that need be specified for voluntary movement are stiffness and resting length: Kinematic variations in displacement, velocity, and acceleration are consequences of the parameters specified, rather than controlled variables, and sensory "feedback"--at least in the conventional computational sense--is not required (cf. Fitch & Turvey, 1978; Kelso, Holt, & Flatt, 1980; Kelso, Holt, Kugler, & Turvey, 1980).

It is worth noting that the view expressed above is equally applicable to disruptions that are static and anticipated (as in the present bite block experiment), and those that are time varying and unanticipated. For example, recent studies of the latter kind have shown that "compensatory responses" of short latency are observed in perturbed articulators as well as in others that contribute to the same "vocal tract goal" (cf. Abbs, 1979, for review). Current theorizing, however, offers two distinct mechanisms to explain the system's reaction to perturbation: A predictive simulation mechanism for anticipated disruptions (Lindblom et al., 1979) and a closed-loop peripheral feedback mechanism for unanticipated disruptions (Abbs, 1979).

The analysis offered here views such a distinction as redundant. Immediate adjustment to either type of perturbation is a predictable outcome of a dynamical system in which muscles function cooperatively as a single unit. If the operation of certain variables is fixed, as in the bite block, or unexpectedly altered, as in online perturbation, linked variables will automatically assume values appropriate to the constraint relation (as long as biomechanical limitations are not violated). In short, dynamical systems (of which speech is a member) always operate in a mode that one can describe as "compensatory."

Although we cannot offer a detailed description of the muscles of the vocal tract in terms of the style of control outlined above, we believe there are some grounds for optimism. Fujimura and Kakita (1979), for example, have performed a three-dimensional simulation of the tongue that uses quantitative control of contractile forces of the muscles actually involved. By treating the tongue muscles (in this case the posterior and anterior portions of genioglossus) as a cooperative unit and maintaining the relative magnitude of contractile inputs to each muscle, it can be shown that the acoustic pattern for the vowel /i/ is obtainable with a wide variety of absolute force values. Thus, as long as the contractile balance among linked muscles is preserved, the exact magnitude of muscle contraction (beyond a critical value) does not matter (see also Kakita & Fujimura, 1977). The generality of this model is limited, at this time, to a single point vowel. Nevertheless, the nonlinear relationship between muscle forces and acoustic pattern allows, or rather provides for, a context-conditioned production system. As in recent accounts of limb localization, invariant "targets" can be attained with different stiffness specifications, as long as the balance in stiffness among relevant muscles is preserved.

As a final point, the analysis offered here suggests a commonality in function between the system capable of producing vowels and that involved in the attainment and maintenance of limb postures. Both systems are materially distinct from each other but share behaviors qualitatively like a nonlinear mass spring. The nontrivial claim, then, is that speech and limb movements are dynamically alike in sharing a common solution to the equifinality problem.

REFERENCES

Abbs, J. Speech motor equivalence: The need for a multi-level control model. _Proceedings of the Ninth International Congress of Phonetic Sciences_, 1979, _2_, 318-324.

Bertalanffy, L. von. _General system theory_. Hammondsworth, England: Penguin University Press, 1973.

Bizzi, E. Central and peripheral mechanisms in motor control. In G. E. Stelmach & J. Requin (Eds.), _Tutorials in motor behavior_. Amsterdam: North-Holland Publishing Company, 1980.

Bizzi, E., Dev, P., Morasso, P., & Polit, A. Effects of load disturbance during centrally initiated movements. _Journal of Neurophysiology_, 1978, _41_, 542-556.

Fel'dman, A. G. Functional tuning of the nervous system with control of movement or maintenance of a steady posture--II. Controllable parameters of the muscles. _Biophysics_, 1966, _11_, 565-578.

Fel'dman, A. G. Superposition of motor programs. I. Rhythmic forearm movements in man. _Neuroscience_, 1980, _5_, 81-90.

Fitch, H., & Turvey, M. T. On the control of activity: Some remarks from an ecological point of view. In R. Christina (Ed.), _Psychology of motor behavior and sports_. Urbana, Ill.: Human Kinetics, 1978.

Fowler, C. _Timing control in speech production._ Bloomington, Ind.: Indiana University Linguistics Club, 1977.

Fowler, C. A., Rubin, P., Remez, R. E., & Turvey, M. T. Implications for speech production of a general theory of action. In B. Butterworth (Ed.), _Language production_. New York: Academic Press, 1980.

Fowler, C. A., & Turvey, M. T. Immediate compensation in bite-block speech. _Phonetica_, 1980, _37_, 306-326.

Fujimura, O., & Kakita, Y. Remarks on quantitative description of the lingual articulation. In B. Lindblom & S. Öhman (Eds.), _Frontiers of speech communication research_. London: Academic Press, 1979.

Gay, T., Lindblom, B., & Lubker, J. Production of bite-block vowels: Acoustic equivalence by selective compensation. _Journal of the Acoustical Society of America_, 1981, _69_, 802-810.

Gay, T. J., & Turvey, M. T. Effects of efferent and afferent interference on speech production: Implications for a generative theory of speech motor control. _Proceedings of the Ninth International Congress of Phonetic Sciences_, 1979, _2_, 344-350.

Gelfand, I. M., Gurfinkel, V. S., Tsetlin, M. L., & Shik, M. L. Some problems in the analysis of movements. In I. M. Gelfand, V. S. Gurfinkel, S. V. Fomin, & M. L. Tsetlin (Eds.), *Models of the structural-functional organization of certain biological systems*. Cambridge, Mass.: MIT Press, 1971.

Hebb, D. O. *The organization of behavior*. New York: Wiley, 1949.

Kakita, Y., & Fujimura, O. A computational model of the tongue: A revised version. *Journal of the Acoustical Society of America*, 1977, *62*, S15. (Abstract)

Kelso, J. A. S. Motor control mechanisms underlying human movement reproduction. *Journal of Experimental Psychology: Human Perception and Performance*, 1977, *3*, 529-543.

Kelso, J. A. S., & Holt, K. G. Exploring a vibratory systems analysis of human movement production. *Journal of Neurophysiology*, 1980, *43*, 1183-1196.

Kelso, J. A. S., Holt, K. G., & Flatt, A. E. The role of proprioception in the perception and control of human movement: Toward a theoretical reassessment. *Perception & Psychophysics*, 1980, *28*, 45-52.

Kelso, J. A. S., Holt, K. G., Kugler, P. N., & Turvey, M. T. On the concept of coordinative structures as dissipative structures: II. Empirical lines of convergence. In G. E. Stelmach (Ed.), *Tutorials in motor behavior*. New York: North-Holland Publishing Co., 1980.

Ladefoged, P., DeClerk, J., Lindau, M., & Papçun, G. An auditory-motor theory of speech production. *UCLA Working Papers in Phonetics*, 1972, *22*, 48-75.

Lindblom, B., Lubker, J., & Gay, T. Formant frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive simulation. *Journal of Phonetics*, 1979, *7*, 147-161.

Lindblom, B., & Sundberg, J. Acoustical consequences of lip, tongue, jaw, and larynx movement. *Journal of the Acoustical Society of America*, 1971, *50*, 1166-1179.

Lubker, J. F. The reorganization times of bite-block vowels. *Phonetica*, 1979, *36*, 273-293.

MacNeilage, P. F. Motor control of serial ordering of speech. *Psychological Review*, 1970, *77*, 182-196.

MacNeilage, P. F. Distinctive properties of speech control. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior*. Amsterdam: North-Holland, 1980.

Perkell, J. S. On the use of sensory feedback: An interpretation of compensatory articulation experiments. *Proceedings of the Ninth International Congress of Phonetic Sciences*, 1979, *2*, 358-364.

Polit, A., & Bizzi, E. Processes controlling arm movements in monkeys. *Science*, 1978, *201*, 1235-1237.

Raibert, M. H. A model for sensorimotor control and learning. *Biological Cybernetics*, 1978, *29*, 29-36.

Thilander, B. Innervation of the temporo-mandibular joint capsule in man. *Trans. Roy. Sch. Dent. Stockholm and Umea*, 1961, *7*, 9-67.

Tuller, B., Harris, K. S., & Gross, B. Electromyographic study of the jaw muscles during speech. *Journal of Phonetics*, 1981, *9*, 175-188.

Turvey, M. T., Shaw, R. E., & Mace, W. Issues in the theory of action: Degrees of freedom, coordinative structures and coalitions. In J. Requin (Ed.), *Attention and performance VII*. Hillsdale, N.J.: Lawrence Erlbaum, 1978.

FOOTNOTE

[1] Indeed it was after observing the bite block performance of our naive subjects under reduced information conditions that this person offered to participate in the present experiment. This was a magnanimous gesture for which we express our gratitude.

THE STREAM OF SPEECH*

Robert E. Remez+ and Philip E. Rubin

Abstract. The use of sinusoidal replicas of speech signals reveals that listeners can perceive speech solely from temporally coherent spectral variation of nonspeech acoustic elements. This sensitivity to coherent change in acoustic stimulation is analogous to the sensitivity to change in configurations of visual stimuli, as detailed by Johansson. The similarities and potential differences between these two kinds of perceptual functions are described.

Studies have shown that the continuously changing stream of speech can be approximated by a kind of acoustic animation, at the theoretical heart of which is an idealization of the human vocal tract as a *resonant horn* (e.g., Chiba & Kajiyama, 1941; Fant, 1956, 1960; Stevens & House, 1955). The details of the acoustics of speech can thereby be explained by noting that the vocal horn can be constricted at different places along its length, that it may be multiply excited, and, especially, that its shape can be changed rapidly. In practical situations, such as speech synthesis, the assumption of the hornlike properties is tacit, presupposed; the synthesizer speaks by the excitation of a lumped-circuit resonator (or its digital equivalent), which is itself approximate to horns of many types, including vocal tracts.

Although the term "speech stream" is often used to refer to the acoustic products of human vocalization, speech has commonly been studied by conceiving this metaphoric stream as an imbrication of more or less isolable elements, such as steady-state or transitional formant patterns, plosive bursts, band-limited noise, and stretches of silence. In our perceptual accounts, then, the exclusive attention to perceptual effects of specific elements in the acoustic pattern has led us to undervalue the coherence of the speech stream. In contrast to theoretical characterizations of the speech stream emphasizing structural continuity (Liberman, 1970), experimenters find it quite agreeable in practice to treat the perceptually relevant acoustic structure as if it consisted of distinct elements. Within this framework, properties of change in the speech signal figure primarily as a problem; from a dynamic array, the listener must somehow extract static elements or cues, perhaps even by means of a specialized decoding device. However, research continually reveals that

---

*Also *Scandinavian Journal of Psychology*, in press.

+Department of Psychology, Barnard College, Columbia University.

Acknowledgment. This research was supported by grants HD 15672 (to R. E. Remez) and HD 01944 (to Haskins Laboratories) from the National Institute of Child Health and Human Development. We gratefully acknowledge the advice of Alvin Liberman, Ignatius Mattingly, David B. Pisoni, Frank Restle, and Michael Studdert-Kennedy.

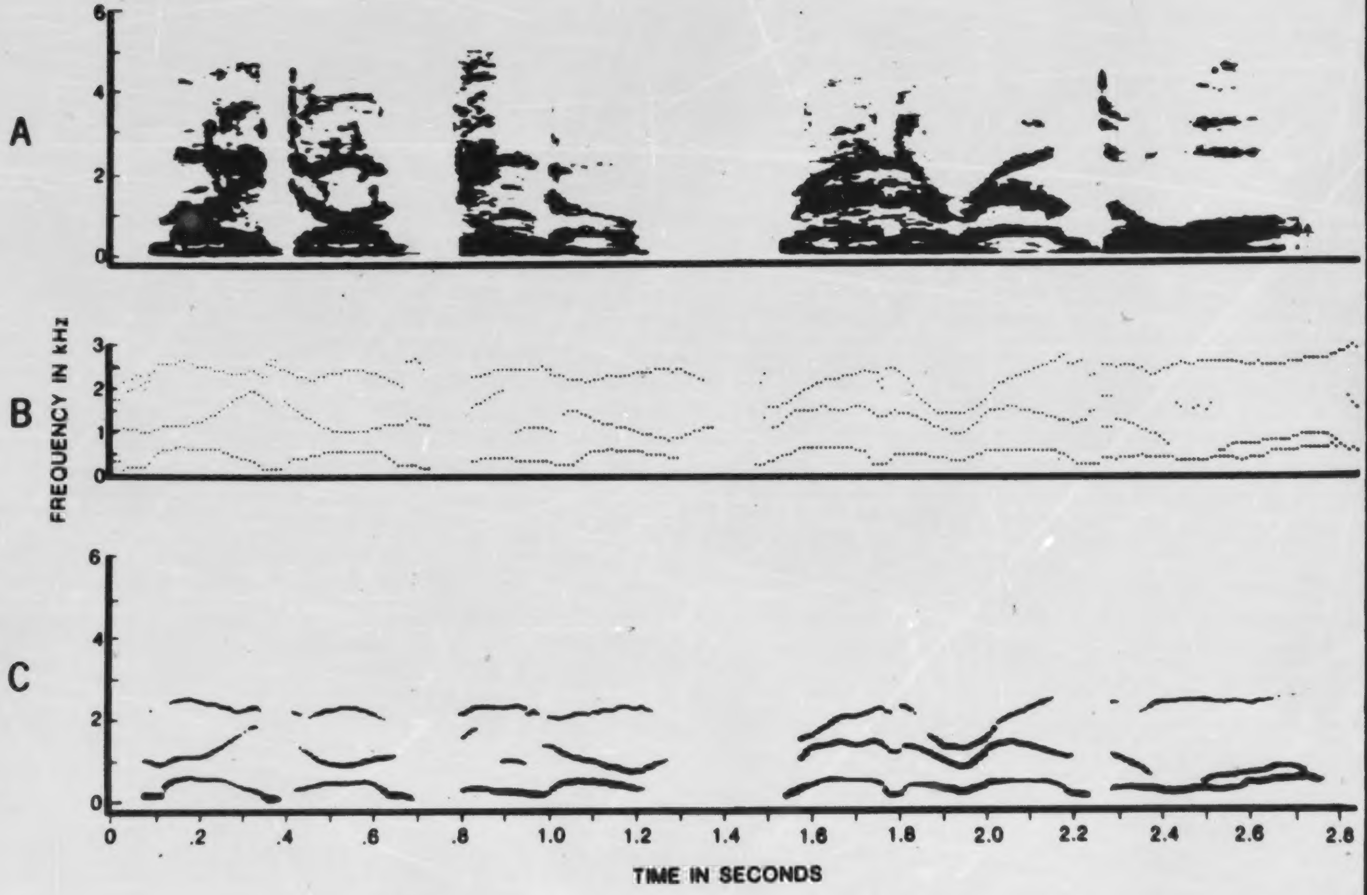


Figure 1. A) The English sentence, "My dog Bingo ran around the wall." The signal was produced by a natural talker and was quantified by a spectrum analyzer employing a bank of acoustic filters. B) The same sentence is represented here as a sequence of center-frequency values of the first three formants. The method of Linear Prediction was used to prepare this series of spectral sections at 10 msec intervals. C) The same sentence is realized here in a signal composed of three time-varying sinusoids. The pattern of relatively continuous change of the natural signal is retained, while the broadband formants and harmonic structure of the natural speech signal are eliminated.

perceivers care very little about the momentary attributes of speech signals. Even under direct test, and in highly favorable conditions, listeners seem unable to report acoustic properties of the stimulation on which their phonetic percepts are reliably based (Best, Morrongiello, & Robson, 1981; Pisoni, 1971; Mattingly, Liberman, Syrdal, & Halwes, 1971).

To describe the speech stimulus in a manner appropriate for perception, we must therefore characterize the time-varying spectrum of the acoustic pattern. In doing so, we elaborate the acoustic coherence of the speech stream, and avoid reducing it to a sequence of static acoustic moments irrelevant to the operating principles of the perceptual system. Only in this fashion may we gain a clue about the "smart" perceptual processes (Runeson, 1977) applied to speech stimuli. In addition, the success of the discrete-cue approximations of speech would then be explained by observing that a sequence of cues, however it reconstitutes properties of coherent change in the speech stream, is perceptible because it conserves the important properties of acoustic variation, rather than because it conserves appropriate short-time spectra or discrete acoustic elements.

RECENT EVIDENCE

Although similar notions have been expressed from time to time in speech research (e.g., Liberman, 1970), several recent experiments especially promote these speculations about the importance of spectrum variation in speech perception (Remez, Rubin, & Carrell, 1981; Remez, Rubin, Pisoni, & Carrell, 1981). In these studies, listeners perceived phonetic segments from acoustic stimuli consisting solely of two or three sinusoids. Frequency and amplitude variations of the sinusoids imitated the changes of the vocal resonances found in natural speech signals (see Figure 1). Specifically, each sinusoidal tone was matched in amplitude and frequency to one of the formants, or resonances, of a natural utterance that served as a model. Matching values were determined for successive 10 msec sections of the natural utterances, and these values were used to control a sinusoidal synthesizer.

None of the acoustic cues typically believed to underlie phonetic perception was present in the sinusoidal patterns: neither formant transitions nor steady state formants, because there were no broadband vocal resonances in this three-tone signal; nor fundamental frequency, because the three-tones were not harmonically related; nor band-limited noise, because the signal had only three periodically unrelated components. Thus, the short-time spectra of the signals did not satisfy the amplitude and frequency requirements of the spectral templates that are sometimes claimed to be useful in analyzing the acoustic pattern into phonetic units (Stevens & Blumstein, 1981). Despite this absence of vocally producible constituents, the overall pattern of frequency and amplitude variation imitated natural acoustic patterns. Listeners who found these sinewave replicas of speech to be intelligible evidently disregarded the inappropriate momentary acoustic structure, and were untroubled by the lack of traditional acoustic cues. Rather, they must have attended to the coherence of the time-variation of the tones, which betrays the vocal origin of the signal, and, at the same time, specifies an impossible sounding voice.

Unlike vocal resonances that share the same laryngeal source, each sinusoid is acoustically independent, and listeners readily reported this distinctness. Accordingly, it was rare that naive listeners spontaneously attended to phonetic information in this grossly unnatural phonetic carrier. Listeners, told nothing in advance about the three-tone signals, heard them simply as three simultaneous tones, modulated asynchronously as if in three-part counterpoint. However, the simple instruction to listen for a sentence enabled almost 70% of naive listeners to detect a sizable chunk of the message, if not its entirety. In other words, listeners made use of phonetic information that was exclusively time-varying in nature, in the absence of short-time spectra characteristic of vocalization.

AN ANALOGY WITH VISUAL EVENT PERCEPTION

The analogy is readily apparent between our experiments on phonetic perception from sinusoids, and Johansson's (e.g., 1975) experiments on the perception of locomotory and other movements from point-light displays. In both cases, it appears that the pattern of coherent change in the stimulation conveys information about the event in progress. In Johansson's case, visual displays are made by videotaping a human figure moving in the dark, illuminated only at the joints of the articulating limbs. Although it is impossible to identify the content of the dot display from the single snapshots, the *moving* dots of light convey a wide range of subtle locomotory information. It is this organized change in the constellation of lights that carries information about the walking actor, despite the absence of static information to reveal which light belongs to which joint. Our sinewave element is like a speech formant in the same way that Johansson's light spot is like a radiocarpal joint--*the value of each element is established only by virtue of the coherent configuration to which it belongs*.

The analogy is not perfect, though. The distal object for Johansson's subject was a walker who seemed to mean nothing by his walking. In contrast, the distal object for our subjects was a message spoken by a strange talker. Our subjects perceived a highly structured message, and Johansson's did not. But, this may merely be a superficial methodological discrepancy if the visual observer can perceive whether the person in the display is performing a tango or a fox trot; or whether the person is using body "language" or American Sign Language, and what the message is (Poizner, Bellugi, & Lutes-Driscoll, 1981). In each case, then, the perceiver identifies a person (one talking, one dancing), a structured transformation (one linguistic, one terpsichoric), and a strange medium (one sinuisoidal, one dotty).

There is an additional discrepancy between Johansson's paradigm and ours. Subjects seem to find the information in the moving dot displays to be more accessible than the information in the sinusoidal displays. We do not understand this very well, but the fact that so many naive subjects hear sinusoids phonetically when instructed to do so may reduce the significance of this difference between the visual and linguistic cases.

A POTENTIAL FORMAL DIFFERENCE

In view of the similarity, it seems appropriate to distinguish formally some properties of visual motion perception and speech perception. Faced with the task of describing the geometry of optical flow, Johansson writes, "the self-motion component in the proximal flow (is combined with) tremendously complex object motion flow. The result is from a mathematical point of view like a chaos or is at least mathematically complex to absurdity" (Note 1, page 6). Of course, the perceiver often disentangles the various components easily despite the limitations we otherwise experience as descriptive geometers. Acoustic change in the case of speech may not prove quite so elusive to describe, though. In principle, the physical limits of the variation of speech sounds are far narrower--constrained anatomically and linguistically--than are the physical limits of optical flow, which appear to be set, after all, by mechanics. And, we have a tremendous head start of forty years of "ecological acoustics" of the vocal tract. In any event, we suggest that studying the _coherence_ of the speech stimulus--describing the _stream of speech_--requires a change of emphasis that brings research on phonetic perception closer to the approach established by Johansson for studying the visual perception of events.

REFERENCE NOTE

1. Johansson, G. _About visual event perception: Perception of motion, dynamics and biological events_. Paper presented at the First International Conference on Event Perception, Storrs, Connecticut, 1981.

REFERENCES

Best, C. T., Morrongiello, B., & Robson, R. Perceptual equivalence of acoustic cues in speech and nonspeech perception. _Perception & Psychophysics_, 1981, _29_, 191-211.

Chiba, T., & Kajiyama, M. _The vowel: Its nature and structure_. Tokyo: Tokyo-Kaiseikan, 1941.

Fant, C. G. M. On the predictability of formant levels and spectrum envelopes from formant frequencies. In M. Halle, H. Lunt, & H. MacLean (Eds.), _For Roman Jakobson_. The Hague: Mouton, 1956, 109-120.

Fant, C. G. M. _The acoustic theory of speech production_. The Hague: Mouton, 1960.

Johansson, G. Visual motion perception. _Scientific American_, 1975, _232_(6), 76-89.

Liberman, A. M. The grammars of speech and language. _Cognitive Psychology_, 1970, _1_, 301-323.

Mattingly, I. G., Liberman, A. M., Syrdal, A. K., & Halwes, T. G. Discrimination in speech and nonspeech modes. _Cognitive Psychology_, 1971, _2_, 131-157.

Pisoni, D. B. On the nature of categorical perception of speech sounds. _Haskins Laboratories Status Report on Speech Research_, Supplement, November, 1971.

Poizner, H., Bellugi, U., & Lutes-Driscoll, V. Perception of American Sign Language in dynamic point-light displays. _Journal of Experimental_

Psychology: Human Perception and Performance, 1981, 7, 430-440.

Remez, R. E., Rubin, P. E., & Carrell, T. D. Phonetic perception of sinusoidal signals: Effects of amplitude variation. Journal of the Acoustical Society of America, 1981, 69, S114.

Remez, R. E., Rubin, P. E., Pisoni, D. B., & Carrell, T. D. Speech perception without traditional speech cues. Science, 1981, 212, 947-950.

Runeson, S. On the possibility of "smart" perceptual mechanisms. Scandinavian Journal of Psychology, 1977, 18, 172-179.

Stevens, K. N., & Blumstein, S. E. The search for invariant acoustic correlates of phonetic features. In P. D. Eimas & J. L. Miller (Eds.), Perspectives in the study of speech. Hillsdale, N.J.: Erlbaum, 1981, 1-38.

Stevens, K. N., & House, A. S. Development of a quantitative description of vowel articulation. Journal of the Acoustical Society of America, 1955, 27, 484-493.

USING THE ACOUSTIC SIGNAL TO MAKE INFERENCES ABOUT PLACE AND DURATION OF TONGUE-PALATE CONTACT

P. J. Price+

Abstract. Productions of /t/ and /d/ in various intervocalic contexts by four native speakers of American English were examined. Closure duration was measured from the acoustic signal and from dynamic palatography data. Place of contact was observed from the palatographic signal, and formant frequency measurements in the closure interval were made. Measurements were averaged over six productions of each stimulus by each talker. Duration measurements by the two techniques correlated well, except (as expected) in instances in which the second vowel had a glottal onset, in which case the duration measurements in the acoustic domain were longer than the durations measured palatographically. Place of contact correlated well with F3 measurements. Normalization across talkers was achieved by dividing each F3 measurement by the talker's mean F2.

INTRODUCTION

Dynamic palatography provides an excellent means to observe tongue-palate contact. This technique is more direct than computing area functions or relying on articulatory introspection. On the other hand, dynamic palatography is less widely available than area function algorithms or introspective data. Dynamic palatography is, thus, of interest both in making direct measurement and in providing a reliability measure for analyses that can be carried out with a wider range of subjects and conditions. The latter is the subject of the present report.

Both closure duration (duration of tongue-palate contact here) and place of constriction are important in the classification of speech sounds. Closure durations can distinguish flap from stop articulations (see, e.g., Port, 1976;

+Now at Massachusetts Institute of Technology, Cambridge, Mass.

Acknowledgment. This research was supported by NICHD grant HD-01994 to Haskins Laboratories. I thank Arthur Abramson, Leigh Lisker, and Michael Studdert-Kennedy for helpful comments on earlier versions. I also thank Tom Baer and Dave Zeichner for their help with the apparatus, and Joe Perkell of M.I.T. for the loan of the RION recording adaptor. An earlier version of this paper was presented at the 101st meeting of the Acoustical Society of America in Ottawa, May 1981 (Journal of the Acoustical Society of America, 1981, 69 (Supplement 1), S56-57). Some of the data also appear in Price (1981).

Zue & Laferrière, 1979), voiced stops from voiceless stops (see Lisker, 1957), and fricatives from affricates (see Dorman, Raphael, & Liberman, 1979). Place of articulation is built more directly into our classification system: Consonants are classified as labial, palatal, or velar depending on where the tongue makes contact with the palate. The artificial palates used permit analysis of articulations from just behind the front teeth to just before the soft palate. Thus, the method is best suited for articulations in the alveolar to the palatal region. Place resolution is to within about 4.5 mm; time resolution is to within 16 msec.

METHOD

Subjects

The study employed four talkers (TB, FB, LL, LR), all native speakers of American English. FB is female; the others are male. FB and TB are practiced talkers with the artificial palate; LL and LR used their artificial palates for the first time in this experiment.

Stimulus Materials

The stimulus times (shown in Figure 1) were read from a randomized list. This list includes items in which [t], [d], and flap [ɾ] occur in intervocalic position. Other items appeared in the list but will not be discussed here. The items are grouped by stress environment and by likelihood of occurrence of flap rather than stop articulations. Flaps can be defined articulatorily by the quick gesture of the tongue tip in the direction of the alveolar ridge associated with them. In American English, flaps typically occur in "latter" and "ladder," but not in "adorn" or "atone." In some contexts, flaps may alternate with stops. These items, therefore, occur twice in the list; they are distinguished by parentheses. Each talker produced six examples of each item.

Measurement Procedures

The artificial palate used in dynamic palatography is about 1.5 mm thick and fits over the hard palate of the talker. It is embedded with 63 gold electrodes configured as in Figures 2a and 2b. Sampling occurs every 15.6 msec. These data can be stored and played back for a frame-by-frame analysis by means of the RION Model DP-01 electropalatograph. The electrodes are about 4 to 5 mm apart, depending on the size of the talker's palate (4.5 mm is used as an estimate here).

Measurements related to closure duration and place of contact were made both acoustically and palatographically. For the purposes of this study, duration and place of closure were defined as follows. Complete closure was defined by contact with at least one electrode in 10 of the 11 columns labeled in Figure 2a. The definition of completeness of closure was flexible since the 15.6 msec sampling interval of the palatograph was somewhat long relative to flap durations (see, e.g., Fisher & Hirsh, 1976; Lisker & Price, 1979; Port, 1976; Zue & Laferrière, 1979), and since the electrodes are tuned as a group rather than individually. Closure duration as measured from the palate,

WORD LIST

TEST ITEMS

FLAPPING ENVIROMENTS

post-stress
heating, heeding
heater, heeder, heat 'er, heed 'er
hotter, solder, water
(Toto), (dodo)
butter
(tódo)
áddict (noun)
potty, toddy

pre-stress
at A, ad A
at Ive

NON-FLAPPING ENVIROMENTS

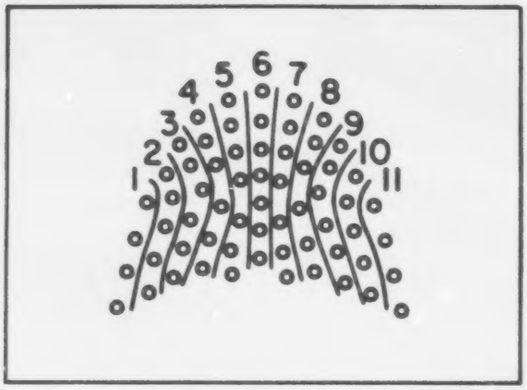
post-stress
tutu (tódo)
(Toto), (dodo)

pre-stress
atone, a tone, adorn
to dó, todó
a dive
addíct (verb)
a day

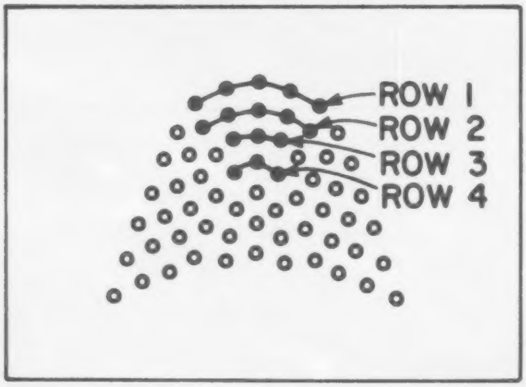
OTHER ITEMS

halter, all day, saunter, center, sender,
party, tardy, hearing, healing, hearer, healer,
horror, holler, sorrow, solo, guru, zulu,
hurry, holly, array, allay, a roan, alone

Figure 1. Word list.


1
2
3
4
5
6
7
8
9
10
11

"COLUMNS" OF THE PALATOGRAPH


ROW 1
ROW 2
ROW 3
ROW 4

"ROWS" OF THE PALATOGRAPH

Figure 2. Configuration of the electropalatograph. a. Columns. b. Rows.

then, was the number of frames of complete closure (as defined above) multiplied by the 15.6 msec frame duration. Closure duration was measured in the acoustic domain by demarcating visually the amplitude dip in the waveform and F1 excitation in the spectral analysis. Place of contact was measured from the palate by taking the mean of the front-most and rear-most rows touched, where "rows" are defined as in Figure 2b. In order to estimate place of contact from the acoustic analysis, frequency measurements of the second, third, and fourth formants were made from spectral cross-sections taken at the end of the acoustically defined closure interval neighboring the stressed vowel.

RESULTS

Duration of Contact

Figure 3 plots the duration measurements made from the palatograph as a function of those made from the acoustic analysis. For items that were pronounced with flaps in all six productions by a given talker or with stops in all six productions, each point represents an average over those six productions by each talker. For items in which flaps alternated with stops in a given lexical item, two points occur, one for the flap articulations and one for the stops. In these cases, the points represent averages over fewer than six productions. The dashed straight line indicates where all points would fall if the two measurements were in perfect agreement.

The triangles in Figure 3 represent the set of tokens judged auditorily to have been produced with a glottal attack on the second vowel. All of these tokens occurred in the V-V́ environment. It is reasonable that acoustic analyses should yield longer closure duration measurements than the palatographic measurements since the low amplitude of the waveform and the lack of strong F1 may result either from closure of the upper vocal tract or from glottal control. The perceptual salience of the glottal attack indicates that the glottis is involved. This figure indicates that some of the flaps may be "glottalized"; note the group of circles with waveform durations of 20 to 40 msec and palate durations of 0 to 20 msec. Very short durations are problematic since the palatograph samples only every 15.6 msec. The averaging over six tokens should, however, moderate this limitation, and, further, some of the differences in the two measurements are larger than the expected maximum difference of about 15 msec. What is more, the amplitude dip in these tokens is somewhat greater than would be expected if upper vocal tract closure were the only mechanism responsible. Thus, it is likely that at least some flaps produced by these talkers involve some glottal control.

There is some indication that the artificial palate may interfere somewhat with articulation (see Hamlet & Stone, 1978). Thus, some caution must be used in interpreting data collected palatographically: Items that sound unnatural may be made with articulations representing more usual habits, while those that sound natural may be made with articulations that reflect adaptive strategies. That the palates used are only 1.5 mm thick minimized but did not eliminate this problem. However, the primary purpose of the present experiment is to map out some acoustic correlates of certain articulations, and for these purposes the limitation is not so crucial as it would be if naturalness were required.

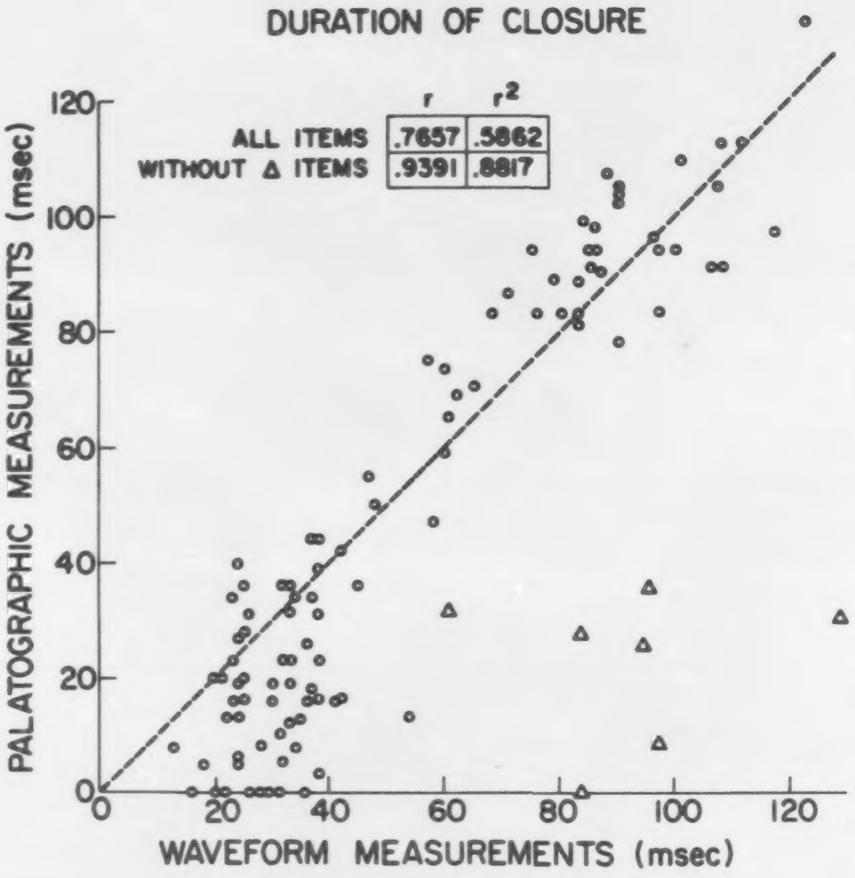


Figure 3. Correlation of duration measurements from waveform with those from palatograph. Each circle is an average over six items produced by a given talker, except in cases where flaps alternate with stops. The dashed line indicates where all points would fall if the two measurements were in perfect agreement. Triangles represent the glottalized items.

In sum, though one must be aware of the limitations of the method, it is possible to make fairly good closure duration estimates from acoustic analyses. This method accounts for 59% of the variance. Further, if the tokens that are perceived as strongly glottalized are removed, 88% of the variance is accounted for.

Place of Contact

The characterization of retroflex sounds by lowering of the third and fourth formants has been noted (Fant, 1973; Ladefoged, 1975). That American English flap has not typically been described as retroflexed may be due more to its phonological role (its alternates with [t] and [d]) than to its perceptual or acoustic qualities. Similarities of American English flap with sounds that are described as retroflex have been pointed out (see Monnot & Freeman, 1972, for a comparison with Spanish single-tap /r/; see Mori, 1929, for a comparison with Japanese flap /r/; see Price, 1981, for a more general discussion). Since the flap productions collected for the present study were acoustically similar to the descriptions in the above-mentioned Ladefoged and Fant references, and since there was corroboration of at least a more posterior place of contact than for [t] or [d], it seemed reasonable to measure F3 and F4 as a potential measure of degree of retroflexion.

The glottalized tokens were omitted in this analysis, since the place of contact and the point at which F3 and F4 measurements were taken do not line up temporally. The stop articulations ([t] or [d]) were also omitted, since the first electrode of the palate was not front enough to measure place of contact accurately: recall that the technique involved the average of front-most and rear-most contact, and stops are produced with a wide area of contact that spreads to the teeth where no electrodes are located. The analysis, thus, concerns the set of all flap articulations.

Figure 4 shows the palatographic place measurements plotted as a function of the third and fourth formant frequency measurements for the practiced talkers (FB and TB). Both F3 and F4 for both talkers correlate rather well with place of articulation as measured from the palatograph: 72-74% of the variance is accounted for in each case. F3 was selected for future measurements, however, since it is easier to track and measure than F4.

These data show that reasonable place estimates can be made for a given talker based on formant frequency measurements. Is it possible, though, to make comparisons across talkers? Absolute frqeuency will not do, for example, since an F3 measurement in the 2000 to 2500 Hz range represents the front-most articulation for talker TB; the rear-most for talker FB. Multiplying these frequency measurements by any constant that represents a given talker's formant frequency range might make normalization possible. Mean F2 was used here and appears to have worked fairly well.

Figure 5 shows how well this normalization worked for all flap productions by all four talkers. Note that the female talker's productions are now within the range of those of the three male talkers. The cluster of points above 1.5 along the x-axis are flaps in the environment of front vowel /i/. Again, the limitation of the placement of the front-most electrode probably makes these articulations appear further back than they are. The correlation

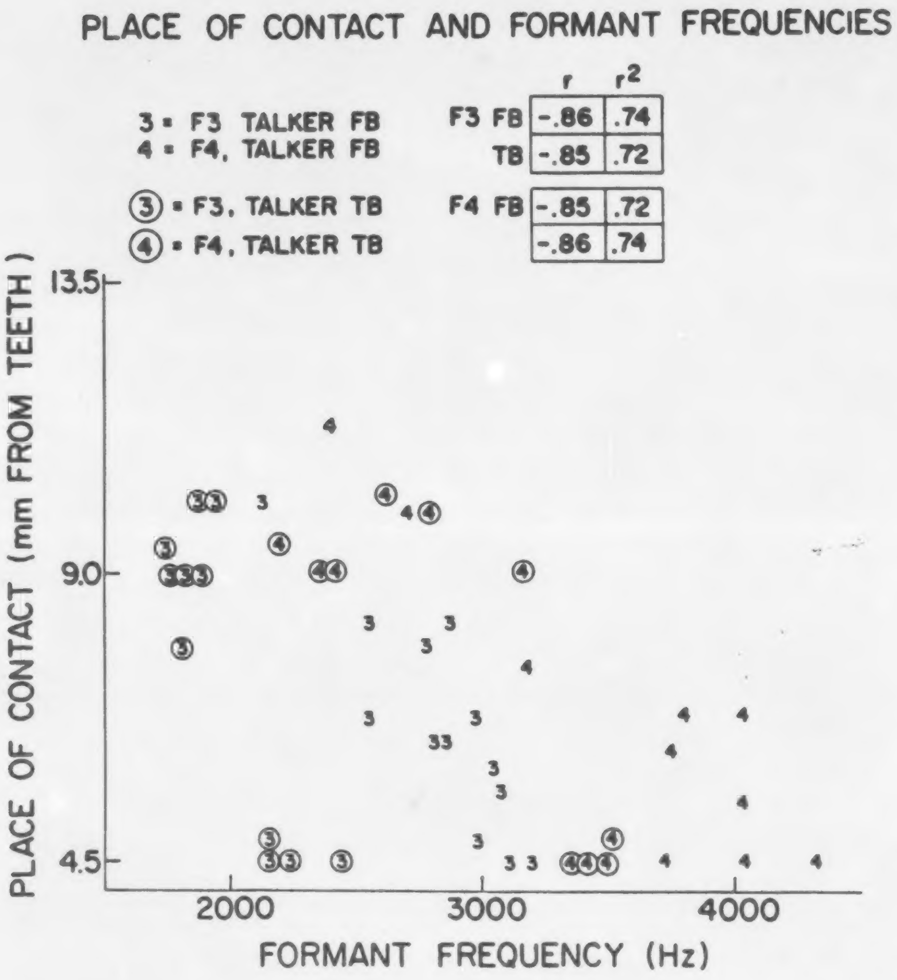


Figure 4. Correlation of third and fourth formant frequencies (F3 and F4) with place of contact for talkers FB and TB.

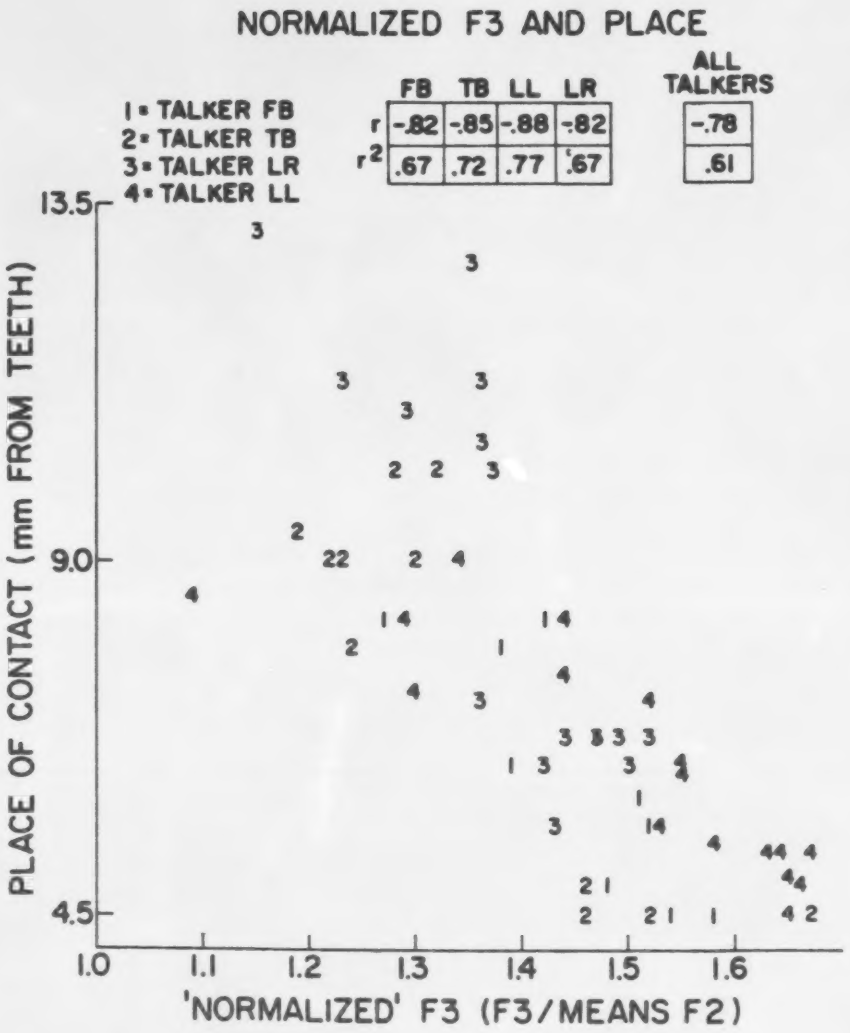


Figure 5. "Normalized" F3 and place of contact. Correlation of F3 divided by each talker's mean F2 with place of contact. All four talkers are plotted here.

coefficients are somewhat higher for the individuals than for the entire group. Yet the group coefficient of -.78 (accounting for 61% of the variance) shows that fairly accurate place estimates can be made from acoustic analysis.

CONCLUSIONS

Subject to the limitations already mentioned, measurements in the acoustic domain can be used to infer both place and duration of tongue-palate contact.

REFERENCES

Dorman, M. F., Raphael, L. J., & Liberman, A. M. Some experiments on the sound of silence in phonetic perception. *Journal of the Acoustical Society of America*, 1979, *65*, 1518-1532.

Fant, G. Descriptive analysis of the acoustic aspects of speech. *Speech sounds and features* (Chapter 2). Cambridge, Mass.: M.I.T. Press, 1973.

Fisher, W., & Hirsh, I. Intervocalic flapping in English. *Chicago Linguistic Society*, 1976, *12*, 183-198.

Hamlet, S. L., & Stone, M. Compensatory alveolar consonant production induced by wearing a dental prosthesis. *Journal of Phonetics*, 1978, *6*, 227-248.

Ladefoged, P. *A course in acoustic phonetics*. New York: Harcourt Brace Jovanovich, 1975.

Lisker L. Closure duration and the voiced-voiceless distinction in English. *Language*, 1957, *33*, 42-49.

Lisker, L., & Price, P. J. Context-determined effects of varying closure duration. In J. J. Wolf & D. H. Klatt (Eds.), *Speech communication papers presented at the 97th meeting of the Acoustical Society of America*. New York: Acoustical Society of America, 1979, 45-48.

Monnot, M., & Freeman, M. A comparison of Spanish single-tap /r/ with American /t/ and /d/ in post-stress intervocalic position. In A. Valdman (Ed.), *Papers to the memory of Pierre Delattre*. The Hague: Mouton, 1972.

Mori, M. *The pronunciation of Japanese*. Tokyo: The Herald-Sha, 1929.

Port, R. F. *The influence of speaking tempo on the duration of stressed vowel and medial stop in English trochee words*. Unpublished doctoral dissertation, University of Connecticut, 1976.

Price, P. *A cross-linguistic study of flaps in Japanese and in American English*. Unpublished doctoral dissertation, University of Pennsylvania, 1981.

Zue, V., & Laferrière, M. Acoustic study of medial /t,d/ in American English. *Journal of the Acoustical Society of America*, 1979, *66*, 1039-1050.

PATTERNS OF HUMAN INTERLIMB COORDINATION EMERGE FROM THE PROPERTIES OF NON-LINEAR LIMIT CYCLE OSCILLATORY PROCESSES: THEORY AND DATA*

J. A. Scott Kelso+, Kenneth G. Holt, Philip Rubin, and Peter N. Kugler

Abstract. The present article represents an initial attempt to offer a principled solution to a fundamental problem of movement identified by Bernstein (1967), namely, how the degrees of freedom of the motor system are regulated. Conventional views of movement control focus on motor programs or closed-loop devices and have little or nothing to say on this matter. As an appropriate conceptual framework we offer the physical theory of homeokinetics of Iberall and his colleagues elaborated for matters of movement by Kugler, Kelso, and Turvey (1980). Homeokinetic theory characterizes biological systems as ensembles of non-linear oscillatory processes, coupled and mutually entrained at all levels of organization. Patterns of interlimb coordination may be predicted from the properties of non-linear limit cycle oscillators. In a set of experiments and formal demonstrations we show that cyclical, two-handed movements maintain fixed amplitude and frequency (a stable limit cycle organization) under the following conditions: (a) when brief and constantly applied load perturbations are imposed on one hand or the other, (b) regardless of the presence or absence of fixed mechanical constraints, and (c) in the face of a range of external driving frequencies from a visual source. In addition, we observe a tight phasic relationship between the hands before and after perturbations (quantified by cross-correlation techniques), a tendency of one limb to entrain the other (mutual entrainment) and that limbs cycling at different frequencies reveal non-arbitrary, sub-harmonic relationships (small integer, subharmonic entrainment). In short, all the above patterns of interlimb coordination fall out of a non-linear oscillatory design. Discussion focuses on the compatibility of these results with past and present neurobiological work, and the theoretical insights into problems of movement offered by homeokinetic physics. Among these are, we think, the beginnings of a principled solution to the degrees of freedom problem, and the tentative claim that coordination and control are emergent conse-

---

*Also in Journal of Motor Behavior, in press.

+Also Departments of Biobehavioral Sciences and Psychology, University of Connecticut, Storrs.

Acknowledgment. We wish to thank Katherine Harris, David Goodman, Betty Tuller, Richard Schmidt, Ronald Marteniuk, and Michael Turvey for their comments on an earlier version of this paper. The research was supported by NIH Grants AM-25814, NS-13617 and Biomedical Research Support Grant RR05596.

quences of dynamical interactions among non-linear, limit cycle oscillatory processes.

1. INTRODUCTION

The beginning of the 1970s brought about a remarkable change in the approach of psychology and related disciplines to the area of movement behavior. This so-called product to process shift that many have remarked on (e.g., Pew, 1974; Schmidt, 1975) evolved from emerging models of human information processing and empirical attempts to discover the nature of a myriad of hypothetical processes--stimulus detection, memory retrieval, response selection (to name but a few)--thought to be involved in behavioral action. Even more significant was the embracing by psychologists of control-theoretic and cybernetic concepts; a move that led to models of motor skill learning (Adams, 1971; Schmidt, 1975) and memory (cf. Laabs, 1973; Stelmach, 1974) and a great deal of laboratory activity (for updates and developments of closed-loop theory, see Adams, 1977; for updates on schema theory, see Shapiro & Schmidt, 1982).

At the beginning of the 1980s it seems timely to remark that our theories and models (like many of the theories and models in biology and the social sciences) are flawed by a deep-seated anthropomorphism that extends back to the time of Descartes: acting humans are compared to machines (computers and servomechanisms) provided with means of control and self-regulation. Motor control theories are peppered with anthropomorphic concepts--perceptual traces, reference mechanisms, comparators, schemas, programs, and so forth--created to "explain" data. Although these concepts have been, and probably will remain, useful for developing an intuitive idea of the way motor systems work, we believe it is now time to consider a dynamical account of movement behavior--one that is consonant with the newly emerging physics of living systems (cf. Kelso, 1981; Kelso, Holt, Kugler, & Turvey, 1980; Kugler, Kelso, & Turvey, 1980, 1982; Turvey, 1980)--even if it is at the expense of some abstraction. The theoretical approach offered here is in its infancy, but the need for it and the types of experiments motivated by it (some of which we report here) can be clarified when we consider further some of the shortcomings of existing theory.

2. SOME LIMITATIONS OF CURRENT THEORY[1]

The problem--aptly understated--that has not received as much theoretical attention in the movement domain as it warrants, is one shared by all living systems; namely, how the internal degrees of freedom of the system are regulated (cf. Bernstein, 1967; Iberall & McCulloch, 1969). Recently we and others have addressed this problem in some detail (e.g., Fowler, Rubin, Remez, & Turvey, 1980; Greene, 1972, 1978; Kelso, 1981; Kelso & Holt, 1980; Kelso et al., 1980; Kugler et al., 1980, 1982; Turvey, 1977, 1980; Turvey, Shaw, & Mace, 1978). One suspects that a main deterrent to a biologically motivated solution to the degrees of freedom problem lies in the view--shared by many--that humans, like computers, are simply information processing devices (cf. Berlinski, 1976). Theoretically, so the argument goes, computers can perform any calculable task that humans perform. Thus it should be possible to retrace the putative sequential steps that a human takes in solving a problem and instantiate them in program form. The physical realization of such a program would exhibit behavior that was artificially intelligent in the sense that such behavior would be indistinguishable from human performance. But, as Berlinski (1976) emphasizes, such a claim rests on a fundamental error. Just because computers can simulate certain types of human performance is hardly evidence that humans actually do employ such programs. Searle (1980) takes a similar stance in noting that the feature that seemed so attractive to people in artificial intelligence, namely the distinction between the program and its realization, proves fatal to the claim that simulation qualifies as duplication. Thus in Searle's (1980) terms:

> The same program could be realized by an electronic machine, Cartesian mental substance or a Hegelian world spirit... If mental operations consist of computational operations on formal symbols, then it follows that they have no interesting connection with the brain [for] the brain just happens to be one of the indefinitely many types of machine capable of instantiating the program. This form of dualism...is Cartesian in the sense that it insists that what is specifically mental about the mind has no intrinsic connection with the actual properties of the brain (pp. 423-424).

A second and related point is that both programming and cybernetical solutions to the degrees of freedom problem vastly undermine the dynamics of the structure to be controlled. They fail, as Yates, Marsh, and Iberall (1972) remark, the "test of matching": in order to couple a control device to the system being controlled there must be some match between scales of energy or mass for efficient operation to ensue. In short, computational or algorithmic solutions place their emphasis on the small signal, information aspects of the system but undervalue the equally important energy-converting machinery (the power fluxes, Yates et al., 1972). A viable account of the dissipation of degrees of freedom for motor systems should recognize the mutuality between informational and power sources.

Much of current theorizing on motor behavior perpetuates a conceptual chasm between the brain as the source of signals for coordination and control and the high power, energy-converting muscular system that is the putative recipient of such messages. For example, in neurophysiological and behavioral studies of movement, it is common for investigators to assume that the products of motor function are isomorphic with the underlying (brain) processes from which those products derived. If the movement of an animal terminates at some spatial location, within a set period of time, for example, that spatial location and duration are said to be contained or represented in the animal's motor program. The motor program then is viewed as causally responsible for generating the spatial position of a limb and metering out the time it takes for the limb to get there.

However, we suspect in real biological systems--unlike formal systems--there may be no need to represent explicitly every detail in the behavioral sequence. Rather, sequential organization may be due primarily to dynamical laws and the existence of constraints that serve to guide those dynamics (cf. Pattee, 1977). If this view is correct, then the order and regularity in movement behavior that we observe will not be due to an a priori prescription--

in terms of programs or reference levels--that is independent of and causally antecedent to the motor activity in question. Nor will it be an isomorphic representation of the behavior to be explained. Rather, spatiotemporal organization--the dissipation of degrees of freedom--will arise as an *a posteriori* fact, an emergent property that is a consequence of, and concomitant with the dynamical behavior of the system (cf. Fowler, 1977).

3. THE DYNAMIC ALTERNATIVE

The answer to what Bellman (1961) called "the curse of dimensionality"--the problem of understanding the relationship between informational and power processes--is offered in a recourse to dynamics, defined as the physics of motion and change. If, as we assume, living systems obey the laws of physics (though they are not readily reducible to them) then, given that formal machine concepts may provide an inadequate basis for complex behavior, what can a dynamical explanation offer in their place? To provide a reasonable answer to this question we have to be armed with certain physical concepts that apply to active living systems.

In the past, a physical description of biological processes has been deemed inappropriate because dynamics has dealt almost exclusively with the behavior of closed, entropic systems (i.e., systems tending towards randomness and disorder). Thermodynamic law states that in a closed system the tendency toward entropy will increase to a maximal value, and that the process is irreversible. In contrast to closed, physical systems, living systems are "open," by virtue of their ability to capture, degrade, and dissipate free energy. As Schroedinger (1945) remarked, living systems "accumulate negentropy" and in so doing maintain their structure and function.

It is only recently that an adequate physics has developed to accommodate the facts of biological systems. Following the lead of Prigogine (1976; Prigogine & Nicolis, 1971) and Katchalsky and Curran (1967), Morowitz (1979) has provided argument that continuous energy flow through a living system constitutes its chief distinguishing feature. In order to prevent the drift towards static equilibrium, biological systems must perform work. Since an isolated, closed system cannot do steady work, it must be connected with a source and a sink; and it is the flow of energy from the source to the sink that constitutes work. *Energy flow, per se, is the chief organizing factor of living systems* (cf. Morowitz, 1979).

All this may seem far removed from a theory of movement, but it leads us to one fundamentally important principle that follows from Morowitz's (1979) main Theorem (p. 33): *the flow of energy through the system from a source to a sink will lead to at least one cycle in the system*. It is the notion that cyclicity provides a dynamic basis for investigating (and understanding) motor systems that we turn our attention to next.

4. CYCLICITY, WITH SPECIAL REFERENCE TO HOMEOKINETIC PHYSICS

(Iberall, 1977, 1978; Iberall, Soodak, & Hassler, 1978;
Soodak & Iberall, 1978; Yates, 1980; Yates & Iberall, 1973).

Persistent cyclicity in biological systems is a non-linear phenomenon; if it were not, the strictures of thermodynamics would ensure a steadily decaying function. Consider, for example, simple mechanical systems such as a mass-spring, in which the equation of motion describes a trajectory towards an equilibrium state. Such systems may be described by a second order differential equation as follows:

$$m\ddot{x} + c\dot{x} + kx = 0 \qquad (1)$$

In Equation 1, oscillatory motion will decay at a rate proportional to the magnitude of the viscous (frictional) term c. This fact is predicated upon the second law of thermodynamics--time flows in the direction of entropy. Yet as we have noted above, living systems are characterized by sustained motion and persistence; they are not statically stable; rather they maintain their form and function by virtue of their dynamic stability. How then, can we ensure sustained motion without violating thermodynamic law?

Consider again the familiar mass-spring equation, but this time with a forcing function $F(\Theta)$:

$$m\ddot{x} + c\dot{x} + kx = F(\Theta) \qquad (2)$$

It is not enough to supply energy to the system described in Equation 2. It must also be supplied at the right place and time in the oscillation. Moreover, the forcing function must exactly offset the energy lost in each cycle for sustained performance to occur (i.e., to satisfy thermodynamic strictures). Many real systems meet this requirement by employing a mechanism--called an escapement--that releases exactly the energy needed to compensate for dissipative losses.[2] The escapement consists of a non-linear element that taps energy from a high potential source--as long as it lasts--to overcome local thermodynamic losses. Thus a pulse or "squirt" of energy is released into the system via the escapement such that, averaged over cycles, the left hand side of Equation 2 equals the right hand side, thereby ensuring sustained motion. Such cycles are called *limit cycles* because they are capable of returning to a stable mode regardless of disturbances that may speed up or slow down the cycle (see below for further details of limit cycle properties).

Real clocks *are* non-linear limit cycle oscillators, that once started, have self-sustaining properties (cf. Andranov & Chaiken, 1949). Many investigators of movement have hypothesized the existence of "clocks" or "metronomes" for purposes of timing (see Keele, 1981, for most recent review) and the rhythmic structure of many biological systems is beyond question (cf. Aschoff, 1979; Oatley & Goodwin, 1971) as is the existence of social-cultural rhythms (e.g., Brazelton, Koslowski, & Main, 1974). However, neither in the field of chronobiology nor in the motor systems area are such hypotheses based firmly in physical theory. Cyclicity (clock-like behavior) in a system arises as a consequence of the transfer of energy from a high potential source to a low potential sink. Cyclicity is ubiquitous in all complex systems, as Yates (1980) has emphatically remarked, because it is an "obligatory manifestation of a universal design principle for autonomous systems."

What does a physically-based theory of periodic phenomena buy us in terms of a principled approach to the coordination and control of movement? The historical antecedent to modern models of the regulation of behavior is the Bernard-Cannon principle of homeostasis. That is, the internal states of an organism are preserved at equilibrium despite changes in the external environment. Modern feedback theories (modeled on quasi-linear servomechanisms) extend the notion of a reference signal-as-goal state to one that can be achieved and adjusted through processes of detection, comparison, and error correction3 (cf. Adams, 1971, 1977; Schmidt, 1975). In sharp contrast, the physical scheme that we outline here is homeokinetic (Iberall, 1970). It is the operating conditions of a configuration of interacting, non-linear limit cycle oscillatory processes that determine the stability (or regulated state) of the organism. According to Iberall and his colleagues (Iberall, 1977, 1978; Iberall et al., 1978; Yates & Iberall, 1973) stability in self-organizing, autonomous systems (the living kind) arises when many cyclical processes become entrained. Thus we can conceive of systemic behavior as being established by an ensemble of non-linear components that are entrained into a coherent configuration.

Elsewhere we have reviewed and presented evidence for the notion--stemming from Bernstein's (1947) initial insights--that a group of muscles, functioning as a unit, exhibits properties qualitatively like that of a non-linear oscillatory system (cf. Asatryan & Fel'dman, 1965; Fowler et al., 1980; Kelso, 1977; Kelso & Holt, 1980; Kelso, Holt, & Flatt, 1980; Kelso et al., 1980, for review). Briefly, we have shown that limb movements may be terminated accurately despite unexpected changes in initial conditions, unpredictable load disturbances during the movement trajectory, functional deafferentation, and all of these in combination. These results have been widely accepted both in animal and human work (e.g., Cooke, 1980; Polit & Bizzi, 1978; Schmidt, 1980) and are interpreted--to a first approximation--as evidence for a mass-spring system (see also Hollerbach, 1980, for expansion of this view to cursive handwriting). But linear mass-spring systems cannot exhibit homeokinetic properties even though they are capable of displaying periodicity. That is to say, the only cycles that meet the non-linear and self-sustaining criteria of biological systems are limit cycles (cf. Goodwin, 1970; Yates, 1980; Yates & Iberall, 1973). A brief discussion of limit cycle properties is thus in order, since it is these that provide a deductive framework for the present experiments.

5. PROPERTIES OF (NON-LINEAR) LIMIT CYCLES

(cf. Minorsky, 1962; Pavlidis, 1973; Sollberger, 1965)

To reiterate briefly, the central feature of homeokinetic physics is the dynamic regulation of a system's internal degrees of freedom by means of coupled ensembles of limit cycle oscillatory processes. In contrast to program and cybernetical conceptions, homeokinetic physics views the existence of active, interacting components and large numbers of degrees of freedom as a necessary and desirable attribute of complex systems. Homeokinetics predicts the discovery of numerous cyclicities and evidence for their interaction. But what is the nature of these cyclicities and what form does their interaction take?

An important caveat at the outset (though it shall not deter us here; instead it is the impetus for the present work) is that the mathematical analysis of non-linear oscillators has hardly begun (cf. Pavlidis, 1973). In contrast, for linear systems, motion in time is relatively easily described, even though the formula describing motion, $x = f(t)$, can be quite complicated. Such functions are conceived of as a system of derivatives, from zero order (the position itself, x) to high order, expressed by a differential equation whose degree is given by the exponential number of the highest differential. In linear systems, motion is regarded as a linearly additive system of first degree differentials whose coefficients may be constants or functions of t, but not functions of x. The family of functions described by linear differential equations are open to solution by various methods of integration.

In sharp contrast, there are no general solutions for non-linear differential equations of motion. For example, in the famous van der Pol equation:

$$\underline{m\ddot{x}} + \underline{k}(\underline{1}-x^2)\,\underline{\dot{x}} + \underline{bx} = 0 \qquad (3)$$

where x = displacement
k = stiffness
b = damping
m = mass

the stiffness coefficient k is itself a function of the dependent variable x, giving rise to the non-linearity and thus negating a unique solution. Fruitful insights into non-linear systems are obtained by graphical methods called phase plane techniques, which plot the first differential, velocity $\dot{x}$ against displacement $\underline{x}$. A set of simple examples is given in Figure 1. Consider a linear differential equation, $\underline{\ddot{x}} + \omega^2 x = 0$. When integrated, the equation yields a set of phase ellipses of the form $(\underline{\dot{x}})^2 + (\omega \underline{x})^2 = c$. One such ellipse is shown in Figure 1A and represents the relation between velocity and position in a simple oscillation. The curves themselves are called phase plane trajectories; it is clear in Figure 1A that the phase plane represents a stable, periodic motion since velocity and position repeatedly return to a certain value (for further details, see figure caption). The spiral trajectory shown in Figure 1B represents an oscillation with continuously decreasing amplitude until it reaches a standstill. A spiral inwards constitutes a damped oscillation; if the direction were outwards (not shown), the oscillation would be unstable with increasing amplitude.

The important point to realize about the phase trajectories illustrated in Figure 1A and 1B is that they are based on linear systems. A change in any parameter, such as the damping coefficient, drastically changes the form of the solution (and thus the phase trajectory). There is then, no preferred solution or set of solutions in a linear system. This is not the case in non-linear systems in which all trajectories tend asymptotically toward a single limit cycle despite quantitative changes in parameter values (see Figure 1C). Thus, a highly important property of limit cycles is their structural

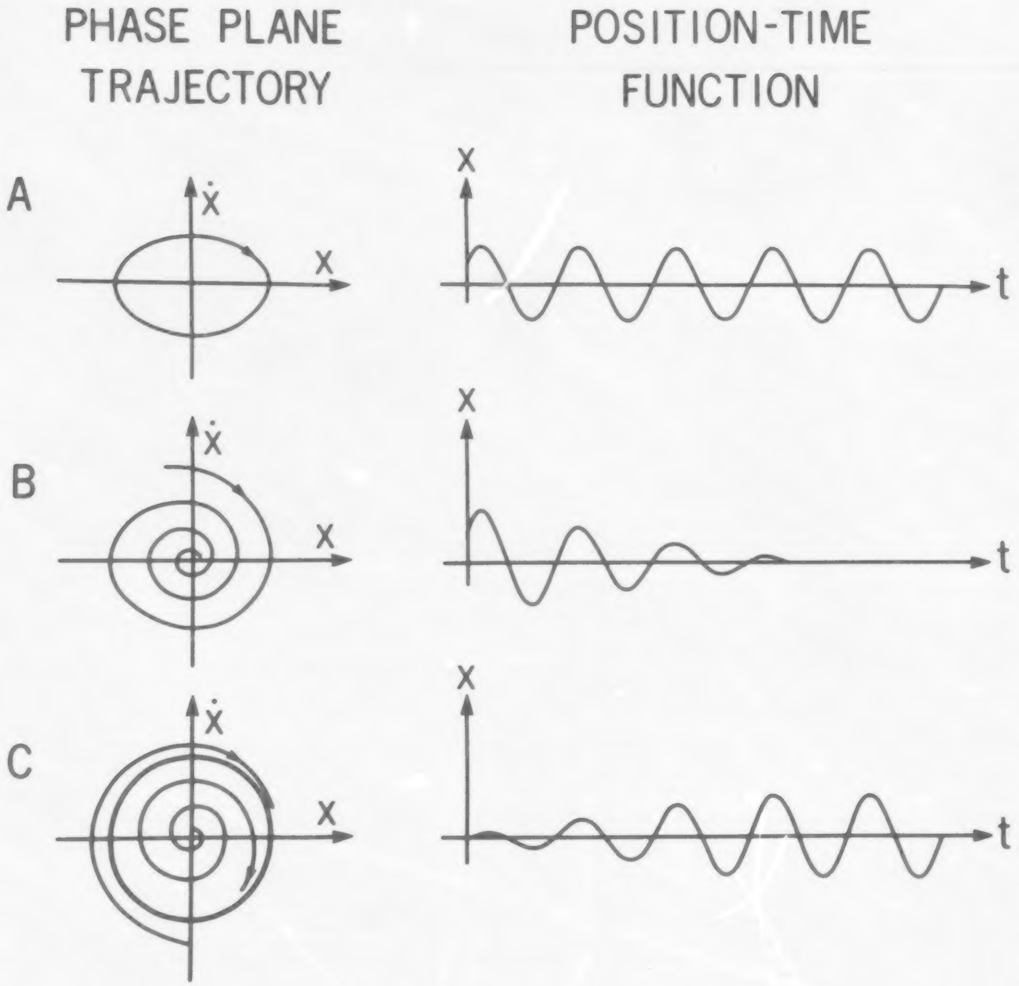

PHASE PLANE TRAJECTORY
POSITION-TIME FUNCTION
A
B
C
$\dot{x}$
x
t

Figure 1. Phase plane trajectories and corresponding position-time functions for three different types of oscillation.
A. Idealized harmonic motion
B. Damped harmonic motion
C. Limit cycle oscillatory motion (see text for details)

stability in the face of variations in parameter values. That is to say, limit cycles exhibit a tendency to maintain a fixed amplitude and frequency (a stable, orbital trajectory) no matter how perturbed (cf. Hanson, 1978; Minorsky, 1962; Oatley & Goodwin, 1971; Pavlidis, 1973). Furthermore, in order for non-linear oscillators to offset precisely the energy lost during each cycle (in the drift towards equilibrium) they must degrade a large amount of free energy (cf. Hanson, 1978; Yates & Iberall, 1973). Because of high energy exchange, non-linear oscillators are quickly resettable following external perturbations. As we shall see, the rapid return of limit cycles to their preferred frequency and amplitude following experimentally imposed perturbations is a predominant feature of the data in the present experiments.

As we have already hinted, and as Pavlidis (1973) further emphasizes, coordination in biological systems arises from cooperative relationships among non-linear oscillator ensembles. In sharp contrast, linear oscillators do not interact among themselves--a fact that is based on the superposition principle in the theory of oscillations (Minorsky, 1962). As defined, the superposition principle says that oscillation in a system of several degrees of freedom consists of a number of component oscillations, each independent of the other.

An essential property of non-linear oscillatory systems, though, is that they always exhibit interaction. Perhaps the chief mode of cooperation among self-sustaining oscillators (and germane to the present experiments) is that of entrainment or synchronization.[4] Apparently, the entrainment phenomenon was first observed by Huygens in the 17th century (cited by Minorsky, 1962). Huygens noted that two clocks whose "ticks" (oscillations) were out of step became synchronized when attached to a thick wooden board. Some two hundred years later, physicists studying electrical circuits and acoustics rediscovered the synchronization effect. When an electrical force of frequency W is applied to an electron-tube oscillator (frequency = W_o), the "beats" of both frequencies are apparent. As the frequencies get closer together the beats diminish until, at a certain difference value $W-W_o$, they disappear entirely and a single frequency, W_1, remains. Similarly when two or more oscillators interact, mutual entrainment occurs (the "magnet effect" of von Holst, 1973) with only a small detuning of their frequencies (Minorsky, 1962). Also, if the frequency of one oscillator is an integer multiple of another to which it is coupled, then subharmonic entrainment--another form of mutual interaction--takes place (also called frequency demultiplication).

In sum, entrainment is an emergent property of a system of non-linear oscillators; it is truly a self-organizing process in the sense that a collection of mutually entrained oscillators functions as a single unit. If indeed biological systems are composed of limit cycle oscillatory processes, then the so-called "degrees of freedom problem"--apparent in much of the current theorizing on movement control--may be minimized through recourse to the entrainment property of non-linear systems. Moreover, entrainment ensures that the degrees of freedom may be dissipated with maximum efficiency and minimum energy cost. Let us consider how this dynamical view interfaces with behavioral experiments.

6. THE PRESENT EXPERIMENTS

There is a rich empirical background in support of the present theoretical perspective. But little of the work is on human subjects (see Cohen, 1970; and Schepelmann, 1979, for exceptions), and none of it is physically (dynamically) founded. Rhythmical phenomena abound at all phylogenetic levels and in many different activities (see Delcomyn, 1980, and Stein, 1976 for reviews). Some of the early work on interlimb phase control by von Holst during the 1930's is classical in this respect, although it is not well known (see von Holst, 1973, for English translation). In his experiments on fish fin movements for example, von Holst identified two main types of coordination. The first of these he termed *absolute coordination*--a 1:1 correspondence between cyclicities of different structures (i.e., where phase and period are the same). The second--much less common interactive state--was called *relative coordination*. Here the fins exhibited different frequencies although at least one of the phases usually corresponded to that observed during the absolute coordination state.

Relative and absolute coordination bear a close similarity to the entrainment properties of non-linear oscillators that we have addressed in the previous section. More recently, Stein (1976, 1977) has elaborated on von Holst's work using the mathematics of coupled oscillators to predict interlimb phase relationships based on the activity patterns of "coordinating neurons" in cockroach and crayfish (see also Graham, 1977). As Stein (1977) notes, the oscillator theoretic approach to neural control is in an embryonic state. But an important first step (which we attempt to take here) is to examine the qualitative predictions of the theory.

For present purposes, there are two empirical antecedents to the following set of experiments. The first is the finding, discussed earlier, that muscles acting at a joint exhibit properties that, to a first approximation, are qualitatively like a mass-spring system. The important point to realize is that such a system is intrinsically rhythmic or cyclic even though it does not have to behave rhythmically or cyclically. Thus, depending on its parameterization, a mass-spring system may or may not oscillate. *Discrete and cyclical behaviors may therefore arise as different manifestations of the same underlying organization* (cf. Fel'dman, 1966; Fowler et al., 1980; Hollerbach, 1980; Kelso & Holt, 1980). The present experiments are continuous with this theme and examine cyclical movements *per se*.

The second antecedent for the current work comes from earlier studies of human interlimb coordination (cf. Kelso, Southard, & Goodman, 1979a, 1979b). When subjects perform movements of the two limbs to different sized targets at different distances from a home position, they do so almost simultaneously. Moreover, the limbs reach peak velocity and peak acceleration at practically the same time during the movements (see also Marteniuk & MacKenzie, 1980). If subjects are required to move both limbs to separate targets, but one limb must hurdle an obstruction on the way, the other unobstructed hand describes a similar arc--at least on the first few trials (Kelso, Putnam, & Goodman, Note 1). When simultaneity is disrupted by instructing the subject to strike one target before the other, there is a sizable cost (either in initiation time or movement errors) compared to temporally compatible, simultaneous conditions (Goodman & Kelso, Note 2; Kelso et al., 1979a; see also Klapp, 1979).

The ubiquity of timing constraints in movement is consistent with the clock-like behavior of mutually entrained oscillators (see Keele, 1981, for further examples, but a different interpretation), which, as we have emphasized, has a firm dynamic basis.5 It remains now for us to examine (and to illustrate) in a more direct manner some of the interlimb relationships predicted by the properties of limit cycles. One of the dominant techniques used to establish the autonomy of rhythms is to perturb the system away from its steady state behavior and observe the manner in which the system reorganizes itself (for example, this procedure is paradigmatic of circadian rhythm research, see Menaker, 1976). Only a system of non-linear oscillators will exhibit maintenance of frequency and amplitude (i.e., a bounded phase trajectory or limit cycle, see Figure 1C) despite perturbations. Thus in the first four experiments reported here (Series A), we examine the foregoing prediction in a simple way. As the fingers of the two hands perform cyclical movements, one or another is unexpectedly perturbed by the injection of a brief torque load or a constantly applied load supplied by D.C. torque motors. Our interest is to define the qualitative features of the system's response to perturbations as revealed in phase, amplitude, and frequency characteristics of each hand.

The remaining aspects of non-linear oscillatory systems that we shall pursue here focus on resonance and entrainment properties. First, non-linear oscillators (unlike linear ones) do not exhibit resonance--an amplitude increase--when driven at their preferred frequency. As noted by Iberall and McCulloch (1969), a fundamental observation about human motor behavior is that externally it does not appear to have a strong metric, but interally it must have. Complex humans operating homeokinetically are self-timed and not tied (in any stimulus-response manner) to external cue constraints. Thus the stability properties of limit cycles predict that their orbit should remain fixed over a fairly wide range of driving frequencies. Second, non-linear oscillators that are coupled together will exhibit (a) mutual entrainment if allowed to function at their preferred frequency and (b) subharmonic (small-integer) entrainment if one is driven at a different rate from the other. As we shall see, the data from the remaining experiments (Series B and C) are a testimony to the powerful entrainment feature of non-linear (limit cycle) oscillators.

GENERAL METHODS AND PROCEDURE

The apparatus used in all the experiments to be reported here consisted of a finger positioning device and associated programming electronics, the details of which are described in an earlier paper (Kelso & Holt, 1980). Essentially the apparatus consisted of two freely rotating supports that allowed flexion and extension of the index fingers about the metacarpophalangeal joint in the horizontal plane. Situated above the center of rotation of each support were programmable D.C. torque motors. An electronics control package permitted programming of torque motor output with respect to movement of the finger in either direction. Thus, in some of the experiments (1 to 4; Series A) perturbations of either digit could be applied for a short duration (termed "brief load") or for a prolonged period ("constant load"). Potentiometers mounted over the axes of motion provided analog signals from which digital representations were obtained via analog-to-digital conversion at a sampling rate of 200 Hz. Additional onset and offset of load perturbation (Series A) and a metronome timing signal (Series C) were specified by analog impulses. A package of conversion routines--implemented on a PDP-11/45 computer--was used to convert the digital signals and display them as time-domain displacement tracings. Additional programs provided mean amplitude, frequency, duration and phase information (cf. Goodman, Rubin, & Kelso, Note 3).

On entering the laboratory, subjects were familiarized with the movement apparatus and seated in a dental chair so that both arms and hands could be comfortably secured in the positioning device. The procedures were described to the subject and any questions answered. The task throughout all experiments was to move the index fingers of one or both hands in a cyclical (flexion-extension) manner from the onset of an auditory start signal to a stop signal. For the two-handed task, the movements were always symmetrical, i.e., flexion (extension) of one hand was accompanied by flexion (extension) of the other. Instructions in all the experiments were to move the finger(s) continuously over a 10 sec trial in a way that felt most comfortable and required least effort.

A. PERTURBATION EXPERIMENTS (1 THROUGH 4)

Specific Methods and Procedures

The subjects were right-handed male and female volunteers who were paid for their services. Two groups of subjects were used--a single-hand group (N=6), that completed the experiments using the index finger of only one hand, and a bimanual group (N=6) that used the index fingers of both the left and right hand simultaneously. In the bimanual group, trials were randomized so that the subject received either no perturbation, a perturbation of the left index finger, or a perturbation of the right index finger. Four trials were given in each of the three conditions, a total of 12 trials for each experiment. In the single-hand group, subjects were told which finger to move and a given trial was either perturbed or not perturbed.6

In Experiments 1 and 2, clamps were placed on the protractor arm of the apparatus, the purpose of which was to constrain movements to a 50 degree range (40-90 degree extension). In Experiment 1, the perturbing torque was set at 50% of maximum torque output available (40.8 oz-in) for a duration of 100 msec, and was introduced approximately midway through the 10 sec trial. The perturbation was injected in a flexion direction as the finger moved through the extension phase of the cycle. In Experiment 2, the torque load was reduced to 25% of maximum (20.4 oz-in) and was introduced about half way through the trial and maintained until completion of the trial. The constant torque load was applied in the direction of limb flexion. Thus, during extension of the index finger the load opposed movement; during finger flexion the load acted in the same direction.

Experiments 3 and 4 were replications of 1 and 2 with the exception that the constraints were removed, allowing subjects to select both amplitude and frequency of movement.

Data Analysis

Measurements of movement amplitude and frequency were obtained from converted digital signals. This procedure relied upon a computer-based determination of displacement maxima and minima per cycle in order to arrive at cycle amplitude. Mean values of amplitude and frequency were obtained only for those cycles before and after the perturbation cycle, and appropriate comparisons made using paired t-tests.

To examine the phasic relationships between the two limbs (in the bimanual group), two analyses were employed. The first involved a computer generated display in which the displacement tracings were superimposed such that any phasic lag or difference between them could be immediately observed. For example, the lower tracing on Figure 2 shows an almost perfect superimposition of the upper individual tracings from each individual hand.

In order to quantify phasic relationships, a second method compared the two traces using cross-correlation techniques. The cross-correlation function for two sets of time-domain data describes the general dependence of the values of one set of data on another. For example, if $u(t)$ corresponds to the signal for the right index finger and $v(t + \tau)$ corresponds to the signal for the left index finger delayed by the interval τ(tau), an estimate of the cross-correlation function for a given tau may be obtained by taking the average product of the two values over the observation time T. The resulting average product will approach an exact cross-correlation as T approaches infinity. In this way the actual phase lag or tau between the waveforms of each hand and the correlation at that lag were calculated for all cycles before and after the perturbation on each trial.

Results

(a) Frequency between and within limbs: The pre- and post-perturbation analysis for the frequency data in the bimanual group across all four experiments is shown in Table 1. Clearly, the results of this analysis point to a very tight periodic relationship between pre- and post-perturbation cycles within each limb. It may be noted that the frequency changes following a brief perturbation range from 0 to .03 cycles per sec. Similarly, the frequency of a limb that subsequently receives a constant load remains virtually unchanged (0 to .05 cycles per sec). None of the pre- versus post differences in frequency attained statistical significance ($p > .05$).

In the single hand experiments, it can be seen again that frequency changes following brief perturbation and the addition of a constant load are small, although they tend to be a little higher than in the bimanual case (.01 to .07 cycles per second, see Table 2). Again, none of the differences attained statistical significance ($p > .05$), nor is there any systematic directional bias in the small frequency changes that are observed.

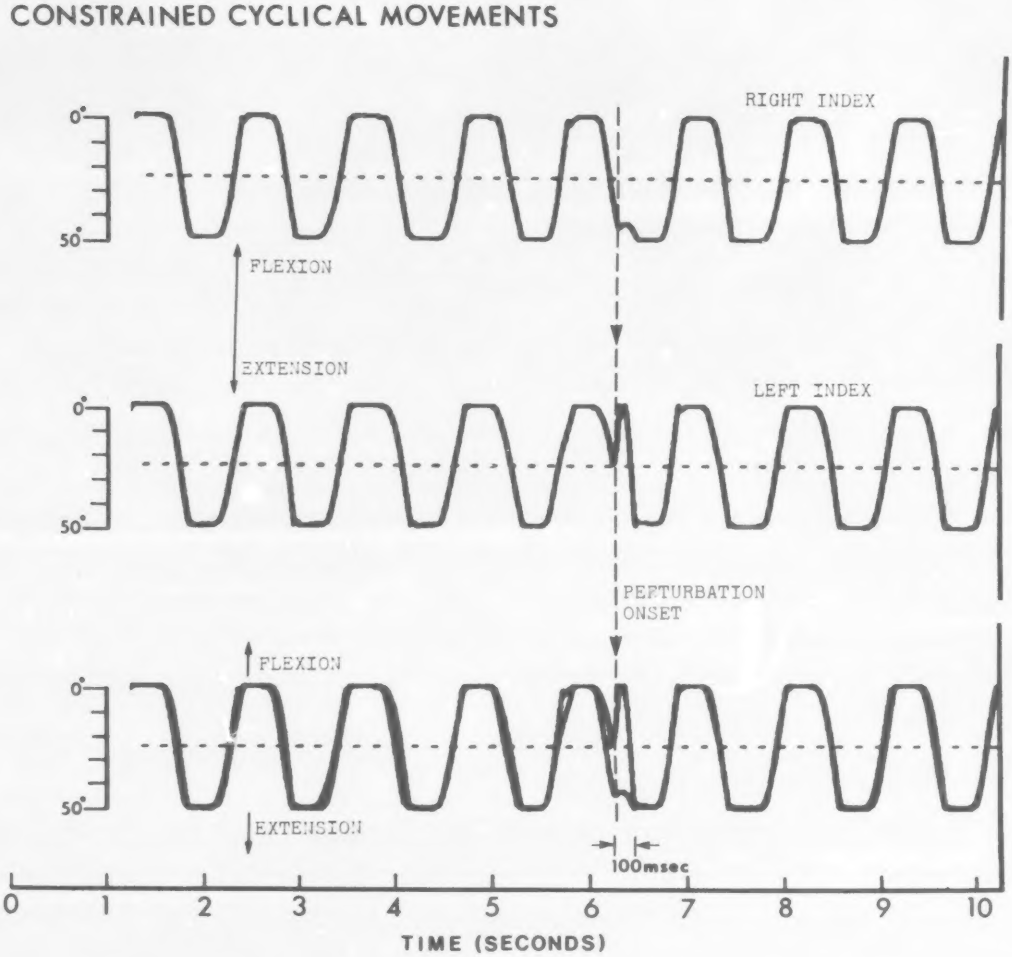


Figure 2. Two-handed cyclical movements between physical constraints. The right index finger was deflected in the direction of flexion as it was extended. In the lower tracing, the two upper tracings are superimposed. The almost perfect overlap reflects the tight period and phase linkage between the fingers. It is interesting to note that the two fingers regained their phasic relationship within the same trial as the perturbation occurred.

Table 1

Means and Standard Deviations of Frequency for Cyclical Movements of Two Hands

			PRE-PERT		POST-PERT		
Experiment 1: Amplitude constrained, brief perturbation	Condition	N[a]	Mean[b]	SD	Mean	SD	t
Right perturbed:	right hand	23	1.39	0.46	1.40	0.44	-0.63
	left hand	23	1.49	0.56	1.46	0.49	0.84
Left perturbed:	right hand	25	1.45	0.50	1.43	0.47	1.98
	left hand	25	1.45	0.47	1.47	0.47	-1.26
Experiment 2: Amplitude constrained, constant perturbation							
Right perturbed:	right hand	22	1.37	0.29	1.34	0.27	1.12
	left hand	22	1.33	0.30	1.28	0.27	1.72
Left perturbed:	right hand	25	1.25	0.31	1.24	0.29	0.58
	left hand	25	1.25	0.29	1.22	0.26	1.73
Experiment 3: Free, brief perturbation							
Right perturbed:	right hand	19	1.68	0.18	1.69	0.16	-0.78
	left hand	19	1.69	0.20	1.71	0.16	-0.53
Left perturbed:	right hand	19	1.72	0.19	1.74	0.18	-1.27
	left hand	19	1.72	0.17	1.73	0.14	-0.29
Experiment 4: Free, constant load							
Right perturbed:	right hand	19	1.79	0.19	1.78	0.13	0.27
	left hand	19	1.79	0.18	1.79	0.13	0.25
Left perturbed:	right hand	19	1.82	0.16	1.84	0.19	-0.68
	left hand	19	1.81	0.16	1.82	0.21	-0.19

[1] Number of data pairs

[b] Frequency in Hertz (cycles per second)

Table 2

Means and Standard Deviations of Frequency for
Cyclical Movements of One Hand

			PRE-PERT		POST-PERT		
a. FREQUENCY[a]	CONDITION	N[b]	Mean	SD	Mean	SD	t
Expt. 1:	Left hand alone	18	1.56	0.43	1.53	0.39	1.52
	Right hand alone	17	1.48	0.38	1.49	0.47	0.35
Expt. 2:	Left hand alone	18	1.57	0.48	1.56	0.48	0.67
	Right hand alone	17	1.58	0.43	1.49	0.31	1.83
Expt. 3:	Left hand alone	19	1.82	0.52	1.77	0.54	2.05
	Right hand alone	18	1.90	0.80	1.88	0.51	0.66
Expt. 4:	Left hand alone	19	1.88	0.57	1.95	0.63	-2.93
	Right hand alone	19	2.01	0.50	1.99	0.55	0.44

[a] In Hertz

[b] Number of data pairs

(b) *Phase relationships between limbs*: The tight periodic relationship in the pre- and post-perturbation cycles within a limb is matched by a tight periodic and phase relationship between the hands. In Figures 2, 3, and 4 examples of the superimposed displacement curves for each experiment are shown. It is clear that despite the imposition of different perturbations, a tight phasic link between the two fingers is preserved. Quantification of these data using cross-correlation techniques verified the extremely close period and phase of the two fingers. The mean tau values and the correlations at those values averaged across trials for each subject are shown in Tables 3 and 4. Absolute values of phase lag (τ) both pre- and post-perturbation are small, and correlations (with only a few exceptions) are extremely high.

(c) *Amplitudes between and within limbs*: Analysis of the amplitude data for the bimanual and single-handed groups is provided in Tables 5 and 6, respectively. In general, amplitudes tended to remain constant with a couple of anomalous findings (see, for example, Experiment 4, Table 5, and Experiments 1 and 2, Table 6).

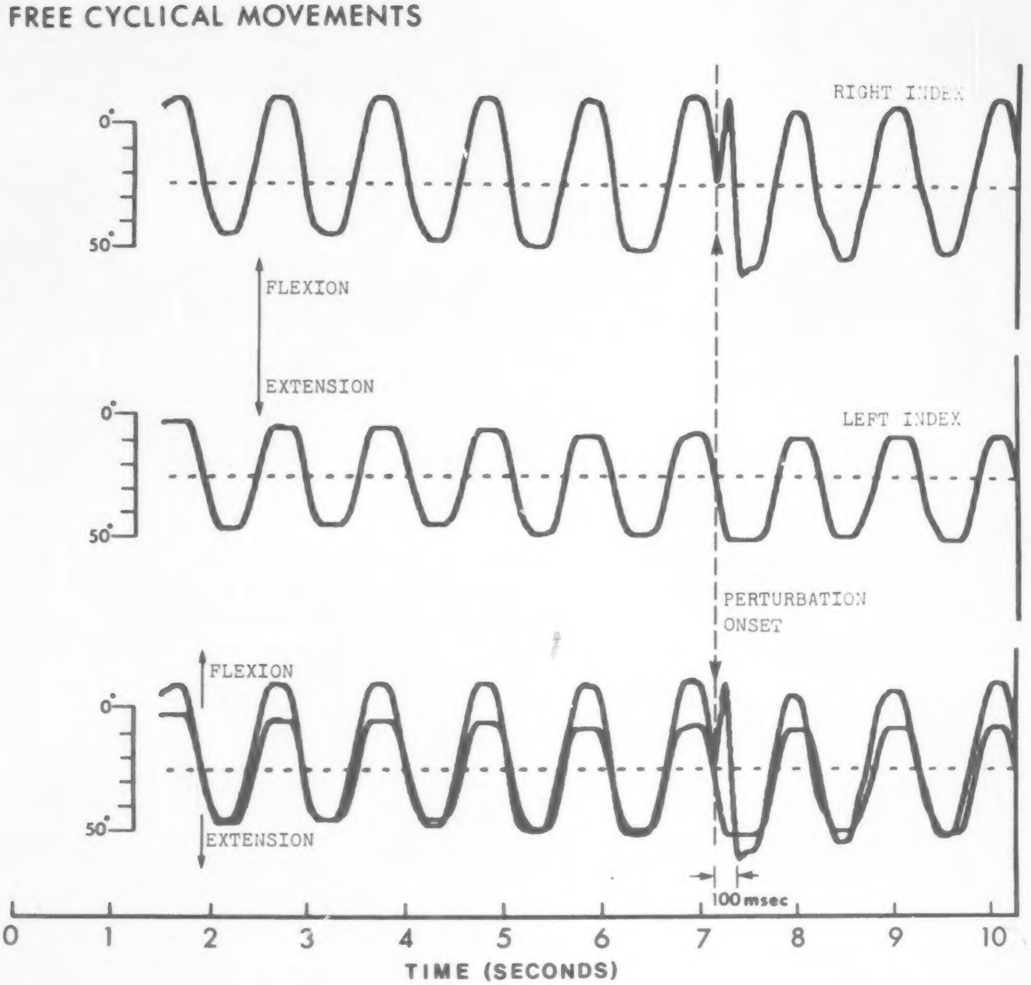


Figure 3. Two-handed cyclical movements without constraints on movement amplitude. Note that the right hand showed a greater range of flexion-extension than the left, with a greater velocity. Phase and period of the two hands are again tightly limited and reentrained within the same cycle after brief perturbation of the right index.

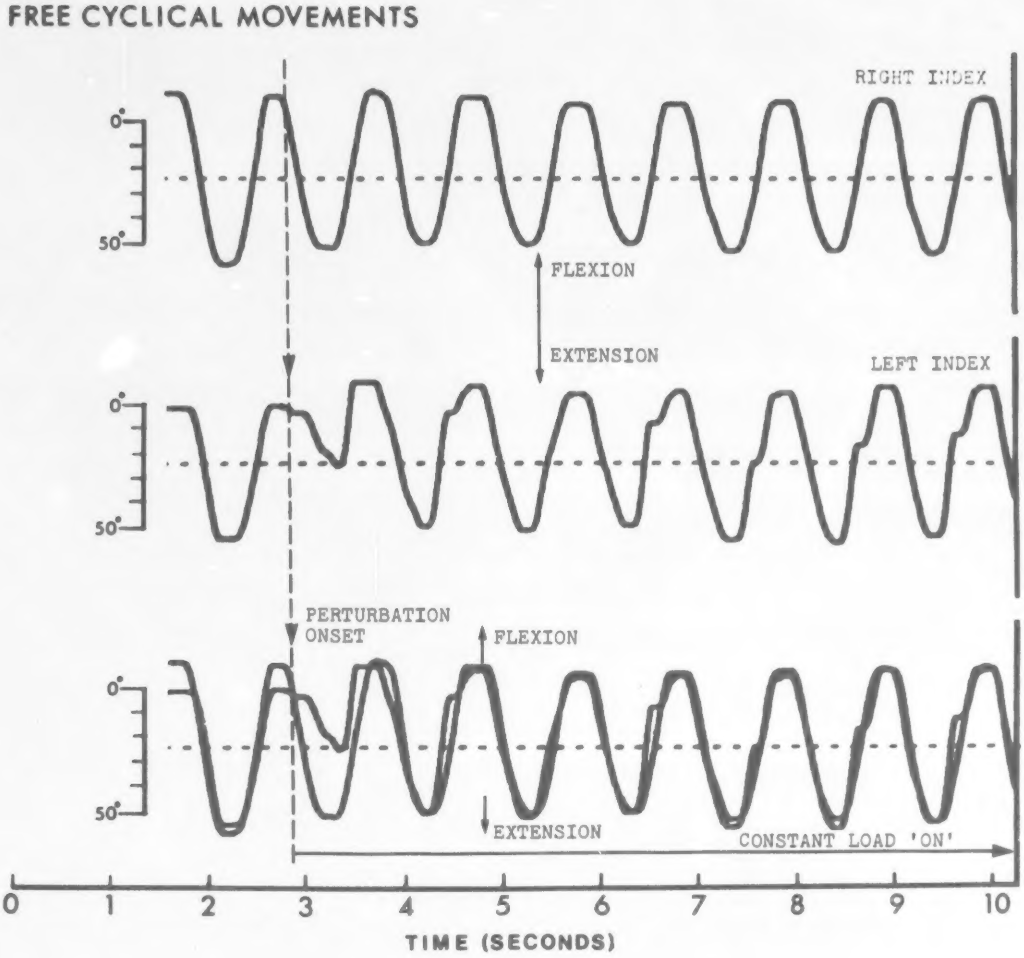


Figure 4. Two-handed cyclical movements without constraints on movement amplitude. A constant load was added to the left index finger in the direction of flexion, while the subject moved in extension. The result on the first cycle was a dramatic decrease in velocity and amplitude. In the remaining cycles, compensation occurred such that not only were phase and period relationships maintained, but also the kinematics--velocity and amplitude.

Table 3

Cross-correlation Analysis of Pre-Perturbation Cycles (Two Hands)

	EXPERIMENT 1				EXPERIMENT 2			
	Correlation		TAU(msec)[a]		Correlation		TAU(msec)	
Subject	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	.94	.17	5.71	3.20	.95	.02	4.29	3.19
2	.97	.02	11.43	4.40	.98	.01	0.00	0.00
3	.99	.01	10.00	9.58	.98	.01	10.63	8.08
4	.98	.01	14.00	4.90	.97	.01	13.75	6.00
5	.97	.01	12.50	8.54	.97	.01	12.50	6.12
6	.97	.04	5.00	5.35	.94	.05	13.75	8.93

	EXPERIMENT 3				EXPERIMENT 4			
	Correlation		TAU(msec)		Correlation		TAU(msec)	
Subject	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	.98	.01	6.25	4.84	.95	.04	15.00	11.02
2	.99	.01	5.80	3.21	.95	.02	17.05	2.50
3	.98	.01	5.63	6.34	.91	.07	15.63	10.74
4	.97	.01	12.50	2.50	.96	.02	13.13	4.29
5	.96	.01	11.43	6.93	.96	.03	11.50	9.27
6	.94	.04	14.38	17.40	.94	.04	16.43	13.02

[a] phase lag with resolution to 5 msec

Table 4

Cross-correlation Analysis of Post-Perturbation Cycles (Two Hands)

	EXPERIMENT 1				EXPERIMENT 2			
	Correlation		TAU(msec)[a]		Correlation		TAU(msec)	
Subject	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	.93	.04	8.57	5.80	.87	.06	4.29	6.39
2	.97	.01	6.43	3.50	*			
3	.99	.01	10.71	4.95	.89	.06	11.25	7.40
4	.98	.00	23.00	9.80	.87	.04	43.50	10.31
5	.98	.02	7.50	5.59	.84	.05	15.63	9.50

	EXPERIMENT 3				EXPERIMENT 4			
	Correlation		TAU(msec)		Correlation		TAU(msec)	
Subject	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	.97	.02	13.75	8.93	.84	.06	10.71	6.78
2	.99	.00	7.36	2.61	*			
3	.98	.01	4.38	4.64	.72	.07	22.50	13.46
4	.96	.17	14.38	7.26	.76	.26	10.00	8.66
5	.96	.02	14.29	11.16	.89	.06	21.88	13.21
6	.90	.09	21.88	20.76	.89	.05	10.29	10.83

*insufficient data

[a] phase lag with resolution to 5 msec

Table 5

Means and Standard Deviations of Amplitude for Cyclical Movements of Two Hands

			PRE-PERT[b]		POST-PERT		
	Condition	N[a]	Mean	SD	Mean	SD	t
Experiment 1: Amplitude constrained, brief perturbation							
Right perturbed:	right hand	23	50.92	2.24	51.50	4.81	-0.78
	left hand	23	50.80	2.16	48.84	5.89	1.78
Left perturbed:	right hand	25	51.16	2.34	50.27	5.47	0.86
	left hand	25	51.93	1.80	51.31	3.02	1.16
Experiment 2: Amplitude constrained, constant perturbation							
Right perturbed:	right hand	22	52.27	3.34	51.27	4.13	1.62
	left hand	22	52.64	6.21	51.24	3.39	1.03
Left perturbed:	right hand	25	51.31	1.77	48.54	8.08	1.69
	left hand	25	51.29	1.81	49.92	4.35	1.52
Experiment 3: Free, brief perturbation							
Right perturbed:	right hand	19	69.76	16.38	67.38	18.27	1.28
	left hand	19	66.29	15.19	64.58	14.35	1.34
Left perturbed:	right hand	19	65.25	14.82	66.35	16.42	-0.98
	left hand	19	70.52	20.70	70.27	21.86	0.20
Experiment 4: Free, constant load							
Right perturbed:	right hand	19	71.77	25.43	75.91	30.49	-2.26*
	left hand	19	67.15	16.50	68.00	13.79	-0.80
Left perturbed:	right hand	19	65.40	18.01	74.05	20.38	-4.44**
	left hand	19	64.50	18.50	60.72	21.28	-1.98

[a] Number of data pairs
[b] Amplitude in degrees
*$p < .05$
**p .01

Table 6

Means and Standard Deviations of Amplitude for
Cyclical Movements of One Hand

b. AMPLITUDE[a]	CONDITION	N	Mean	SD	Mean	SD	t
Expt. 1:	Left hand alone	18	50.56	2.73	50.71	2.86	-1.53
	Right hand alone	17	55.05	3.70	56.77	5.85	-2.88*
Expt. 2:	Left hand alone	18	51.68	2.71	52.25	2.35	-3.35*
	Right hand alone	17	57.96	4.71	58.80	5.70	-1.15
Expt. 3:	Left hand alone	19	60.13	24.33	61.15	23.63	-0.81
	Right hand alone	18	63.10	25.29	60.49	21.94	1.59
Expt. 4:	Left hand alone	19	58.43	22.75	60.35	13.53	-0.60
	Right hand alone	19	53.88	12.95	51.46	13.66	1.51

[a] in degrees

*p < .05

However, in several of the analyses of free bimanual movements (unconstrained), it is clear that the amplitude differences almost reached significance and, given greater power, may have done so. Of course the foregoing analysis is quite global in the sense that it clouds potential amplitude adjustments that may have occurred on individual cycles before or after perturbations. With this in mind, we compared amplitudes of individual cycles in Experiments 3 and 4 (single and bimanual groups) on those trials in which at least five cycles preceded and followed the perturbed cycle (P). Table 7 provides a summary of those cycles reaching significance ($p < .05$). Out of a possible number of 28 significant differences within each cell, it can be seen that the maximum number to reach significance is only eight. Furthermore, there does not appear to be any systematic pattern among those cycles that is statistically different from any other. Thus, if subjects were adjusting amplitude over several cycles following the perturbation, we might have expected a larger number of differences between those cycles immediately preceding perturbation onset (P-1, P-2) and those cycles immediately following the perturbation (P+1, P+2). Although there was a greater number of pre- post differences in amplitude overall, there was nothing to suggest--even in the individual amplitude data--any progressive recalibration on the first or second post-perturbation cycle back to pre-perturbation values.

Table 7

t-test Matrix of Total Number of Significant Differences Between Individual Cycles. All Subjects on Expts. 3, 4 One and Two Hands.

Comparison pre-pert cycles

	P_{-5}	P_{-4}	P_{-3}	P_{-2}	P_{-1}[a]
P_{-5}		3	2	1	0
P_{-4}			1	0	2
P_{-3}				1	1
P_{-2}					0

Comparison post-pert cycles

	P_{+5}	P_{+4}	P_{+3}	P_{+2}	P_{+1}[b]
P_{+5}		0	3	2	2
P_{+4}			2	3	3
P_{+3}				4	3
P_{+2}					4

Comparison pre-post

	P_{+1}	P_{+2}	P_{+3}	P_{+4}	P_{+5}
P_{-5}	3	8	5	5	3
P_{-4}	5	4	4	3	2
P_{-3}	3	5	5	3	2
P_{-2}	2	5	6	4	1
P_{-1}	4	6	4	5	4

[a] P_{-N} = N^{th} cycle preceding perturbation

[b] P_{+N} = N^{th} cycle following perturbation

[c] Total possible significant differences in each cell = 28

Discussion

The finding that frequency (period) within a finger remains constant despite the addition of brief and constant loads strongly supports the view that the limbs behave in a manner qualitatively like limit cycle oscillators. The first criterion for limit cycle processes--that frequency (period) and amplitude tend to be maintained despite perturbations--receives good support in our data. Further support for the limit cycle account is evident in the tight phase and period relationships between the fingers of the two hands.

Perhaps the crucial test of the entrainment property of non-linear oscillators is the perturbation experiment. It is clear in all four experiments reported here that, quickly after the perturbation of one finger, both fingers become re-entrained within one or two cycles; there is neither a phase lag nor a difference in periodicity between the two fingers. This result is reminiscent of work in animal locomotion. Thus Shik and Orlovskii (1965) temporarily prevented one of the limbs of a locomoting cat from initiating the transfer phase of its step cycle--a procedure that necessarily disrupted the phase relations among all four limbs. Within a few full cycles, however, the limbs returned to their previously established phase relationships, as in the present experiments.

The general finding that amplitude tends to be preserved in both limbs in spite of load perturbations applied to one, provides additional support for the limit cycle oscillator view and extends previous work on single trajectory movements (e.g., Fel'dman, 1966a, 1966b; Kelso & Holt, 1980; Polit & Bizzi, 1978) to voluntary cyclical movements (see also Fel'dman, 1980). A noteworthy feature of the data is the kinematic differences within and between limbs when the system is perturbed. For example, in Figure 4, even though frequency is maintained in the constant load condition, velocity (as reflected in slope differences) and amplitude differ for the two limbs. This finding draws our attention to a fundamental point that we have made in earlier papers; observed kinematic details are consequences of the system's dynamics (e.g., mass, stiffness, damping) and are determined by those dynamics (cf. Fowler et al., 1980; Kelso & Holt, 1980).

B. FURTHER DEMONSTRATIONS OF ENTRAINMENT

There are two additional properties of non-linear, limit cycle oscillators that we shall consider under the heading of entrainment. The first is that when coupled, oscillators of slightly different frequencies will tend to entrain at some intermediate frequency. As von Holst (1937/1973) remarked, the striking feature of coordinated movements is their "accordance in tempo." He called this frequency detuning or mutual entrainment property "the magnet effect": simply a tendency of one rhythm to impose its tempo on another. A second form of mutual interaction among oscillators occurs if the frequency of one is an integer multiple of another to which it is coupled, a property termed subharmonic entrainment or frequency demultiplication. In the following experimental demonstrations, both types of oscillatory interaction, mutual and subharmonic entrainment, are clearly evident.

1. Mutual Entrainment

The basic tack on this issue was first to determine the preferred frequency of each limb in isolation and then examine possible interactions between the limbs when they perform together. The procedure was similar to that employed in the previous experiments. Each of six, right-handed subjects (none of whom participated in any of the previous studies) completed four trials with the left hand only, the right hand only and both hands combined. The twelve trials, each lasting 10 sec, were randomized for all subjects. The subject was instructed to move the finger(s) cyclically at a frequency and amplitude that felt most comfortable. No constraints or perturbations were imposed at any time during the trial.

Mean amplitude and frequency data for each trial were again obtained from the converted digital signals. The means for all subjects are shown in Table 8. The frequency data meet the predictions of limit cycle oscillators in that the left hand is "attracted" to the right hand, which in turn shows only a very small and statistically insignificant frequency modulation. Clearly the overall effect is modest, as one would expect in an experiment using simple cyclical movements of the two hands. There are obvious ways to amplify the extent to which one limb imposes its rhythm on another by changing, for example, limb dynamics (e.g., mass, lever arm) or by fatiguing one limb or the other. Our intent here, however, has been simply to demonstrate the mutual entrainment effect. That the right hand does exert an "attracting force" on the left in right-handed subjects illustrates another mode of cooperation among oscillators. This finding is also a testimony to the difficulty subjects have in performing two different rhythms at the same time. Entrainment, we suspect, represents a major limitation on what activities can actually be performed. As von Holst (1937/1973, 1939/1973) remarked, it appears to be an important principle of central order. Yet entrainment has received little or no attention in theories of how skills are acquired, or how movements are controlled.

Table 8

Means and Standard Deviations of Preferred Frequency and Amplitudes in One-handed and Two-handed Cases (N=6)

		One hand		Two hands		
		Left	Right	Left	Right	Combined
FREQUENCY[a]	Mean	1.973	2.001	1.996	1.995	1.996
	SD	0.03	0.05	0.03	0.03	0.03
AMPLITUDE[b]	Mean	62.93	66.91	57.57	65.84	61.71
	SD	4.31	3.59	9.17	3.84	4.01

[a] in Hertz
[b] in degrees

2. Subharmonic Entrainment

There is at least one situation in which individuals do not experience difficulty in performing separate rhythms with the hands, and that is when they both share a common time base. Recently Klapp (1979) has shown "interference" between the two hands in a repetitive key pressing task when the periodicity of each is different, as measured by the subject's ability to match a pacing tone. When the periodicities share common timing, even though shifted in phase, no such interference occurs. These results may be accounted for by another property of oscillator interaction, namely, low integer subharmonic entrainment. This form of entrainment is such an overwhelming phenomenon in human motor activity that we simply illustrate it here (see Figure 5). Subjects were asked to move one finger at their preferred frequency and to move the other finger at a different frequency. We show two examples of one subject's performance. As illustrated in Figure 5, the oscillation of the limb moving at slower frequency exactly coincides with the appropriate oscillation at the faster frequency—one is a simple (2:1, 3:2) subharmonic of the other.7 The example also illustrates the interesting property of amplitude modulation (von Holst's superimposition effect). Thus, on some coinciding cycles, a "beat" phenomenon can be observed (particularly in the 2:1 ratio) in which the amplitude of the higher frequency hand increases on the occasion that the lower frequency oscillation takes place. Note also that in the lower frequency oscillation the finger is never completely still. Although its amplitude is much smaller, it is clear that there is a small oscillation, especially in the 2:1 ratio condition. In effect, both fingers are cycling at the same frequency—only the amplitude or force distribution to each finger varies. The foregoing result fits nicely with recent work by Shaffer (1980) on highly skilled pianists. Shaffer showed that the pianist's right hand carrying the melody plays with more "weight" than the left, and that gradual and sudden changes in both hands can be made without disrupting timing. It is tempting to suppose that Shaffer's pianist is displaying (admittedly in a more refined way) the basic superimposition principle for combining the outputs of coupled oscillators (see also Gallistel, 1980, and Kelso, 1981, for additional examples).

It is worth noting that the entrainment properties demonstrated above are not restricted to movements of the fingers but are also apparent in systems that share little or no anatomical similarity. In an analysis of the interrelationships between speaking and manual activity (Kelso, Tuller, & Harris, in press), we have shown that subjects, when asked to repeat a simple syllable (the word "stock") at a different rate from their preferred finger rate, do so by employing low integer subharmonics. The situation is reversed (though not necessarily symmetrically) when subjects are instructed to move their hands at a rate different from their preferred speaking rate. Again, the ratios chosen are always simple ones (e.g., 2:1). We interpret these preferred relationships as emergent characteristics of a non-linear ocillator ensemble; the collection of entrained oscillators functions in a single unitary way. Entrainment therefore ensures a stable resolution of simultaneous temporal processes throughout the whole system. Moreover, the form of entrainment is limited to a relatively restricted range of preferred relationsnips--a feature captured in Iberall and McCulloch's (1969) phrase as an "orbital constellation."

102

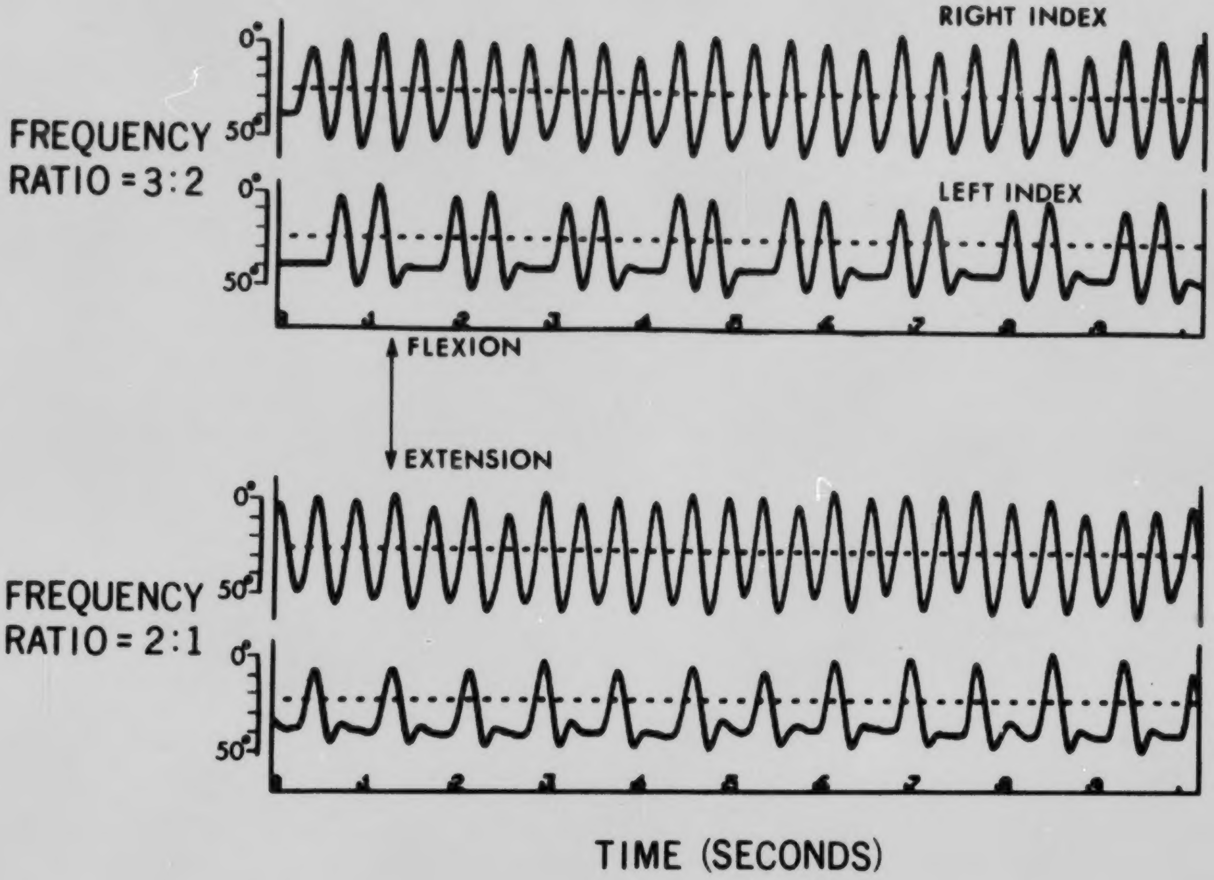


Figure 5. An example of a subject's response to instructions to move the fingers at different frequencies. Beats in the lower frequency hand tended to coincide with the beats of the higher frequency hand. On some coinciding cycles, a "beat" phenomenon can be observed in which the amplitude of the higher frequency hand increases in relation to non-coincident cycles (see especially the 2:1 ratio). In addition, slight movement during the "quiet" phase of the low frequency rhythm is indicative of the highly constrained nature of the two-hand linkage.

C. NON-RESONANCE PROPERTIES

A characteristic of a linear oscillator, but not of the non-linear limit cycle oscillator under examination here, is that when driven at its fundamental frequency it will display resonance--a behavior that results in an increase in amplitude of oscillation. A final experiment examined whether amplitude changes would be observed when the limb was driven by an external rhythmic source at its preferred frequency, and at other frequencies higher and lower than the preferred frequency.

Four paid volunteers, none of whom had participated in the earlier work, took part in this study. The procedures were very similar to those already discussed, with the following exceptions. Before the experiment proper, the preferred frequency for each individual subject was obtained by finding the mean and standard deviation of five trials (10 sec each) in which the subject chose his/her rate and amplitude. Our prevous work indicated that this was more than enough time to achieve stable measures. These data were then used as the basis for driving frequency manipulations effected via a visual metronome. Seven frequency conditions were used (F_o, the subject's preferred frequency $F_o \pm 2$ *SD*, $F_o \pm 4$ *SD*, $F_o \pm 6$ *SD*). For example, if a subject had a measured preferred frequency of 1.5 Hz with a standard deviation of 0.075 Hz, he/she would be asked to produce one flexion-extension cycle to each metronome beat under the following conditions: 1.5 Hz, 1.5 ± .15 Hz., 1.5 ± .30 Hz., 1.5 ± .45 Hz. Five trials were given in each of the seven conditions that were randomized for each subject, and movements were two-handed in all cases.

The mean amplitude and frequency are displayed in Table 9 for all seven driving frequency conditions. These data were analyzed in a 2 x 7 x 5 (hands [left/right] x driving frequency x trials) within subjects analysis of variance. None of the amplitude comparisons was statistically significant, but there were definite effects on frequency in response to the driving stimulus conditions, $F(6, 18) = 19.28$, $p < .01$. None of the subjects had any difficulty performing the task in any of the driving conditions, as confirmed by scanning the graphical output of metronome and displacement waveforms. One such example illustrating a fixed maintenance of amplitude across the most extreme driving conditions is shown in Figure 6.

Nevertheless it is apparent from the mean data that there are tendencies for amplitude and frequency to be linearly related particularly at faster driving frequencies. Paradoxically, and if we were dealing with linear oscillators, amplitude should *increase* with slower driving frequencies--a prediction not borne out by the present data.

In short, the responses of the limbs to different driving frequencies seem to display both linear and non-linear characteristics. This is not particularly surprising for non-linear systems are capable of exhibiting linear behavior over a range of parameter values. The distinguishing feature of non-linear oscillators is that their behavior can be dramatically altered in terms of phase and amplitude when driven at certain frequencies. Much more needs to be done to determine if, and under what conditions, amplitude (and/or phase) changes occur as "saltatory jumps" at certain frequencies; that is, when one stable orbit is forsaken for another.[8]

Table 9

Means and Standard Deviations of Frequency and Amplitude of Two Hands When Driven by Light Pulses at Different Frequencies (N=6)

			FO[c]	FO+2 SD	FO-2 SD	FO+4 SD	FO-4 SD	FO+6 SD	FO-6 SD
FREQUENCY[a]	Left	Mean	1.82	1.92	1.74	2.11	1.61	2.34	1.53
		SD	0.02	0.02	0.03	0.02	0.03	0.03	0.06
	Right	Mean	1.82	1.92	1.74	2.11	1.61	2.33	1.55
		SD	0.03	0.02	0.04	0.03	0.04	0.03	0.04
AMPLITUDE[b]	Left	Mean	44.69	48.40	47.06	40.29	40.85	37.28	44.17
		SD	3.65	5.03	3.64	1.75	3.35	2.54	4.43
	Right	Mean	52.38	49.98	50.84	45.55	46.25	40.10	49.71
		SD	3.01	5.86	3.18	2.41	3.04	1.66	4.44

[a] in Hertz

[b] in degrees

[c] reflects mean preferred frequency

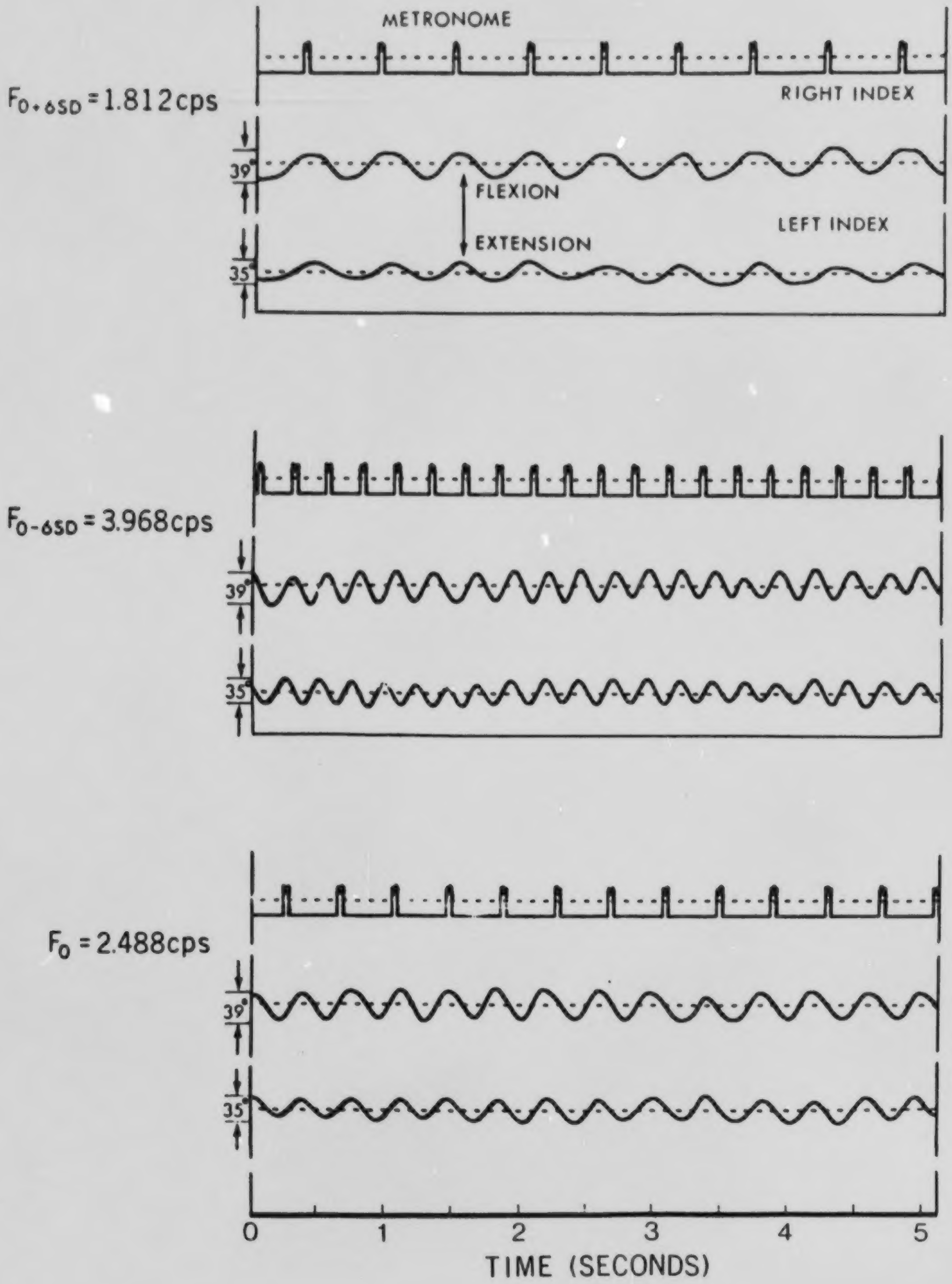


Figure 6. Tracings indicate a constant amplitude in the two hands when driven at different frequencies by a visual metronome (a pulsing LED). The onset of a light pulse in most cases was synchronized with the point of maximum extension or flexion. Slight deviations of the finger oscillations from the metronome period can be observed in some cycles.

The point that should not be lost here is that oscillatory systems can behave in a linear manner yet still be non-linear in nature. Consider the intuitive example of dolphin locomotion. At low speeds, dolphins cruise with little energy cost because of their streamlined shape; water flow over their body surface is laminar. At higher swimming speeds, however, the creation of turbulence can increase energy costs by an order of magnitude (see Brookhart & Stein, 1980). As the relationship between velocity and energy costs becomes non-linear, the dolphin divorces the swimming mode for its novel and more economical "running" mode. In short, scale changes reveal the non-linearities in the system and are capable of effecting qualitative changes in behavior. A major enterprise then for a science of movement becomes one of identifying the necessary conditions under which such bifurcations or phase transitions occur (for clues, see Yates, in press; Yates et al., 1972).

GENERAL DISCUSSION

In previous papers and in the introduction to this one, we have considered ways in which the motor system might solve the degrees of freedom problem, first identified by Bernstein (1967). Our approach is double-edged: first, we have sought to identify and analyze rigorously functional groupings of muscles--synergies or coordinative structures--constrained to act in a unitary way; and second, we have attempted to establish the language of dynamics as the appropriate vocabulary upon which to rationalize the properties of coordinative structures. More specifically, we have provided theoretical grounds suggesting that a group of muscles organized as a single functional unit possesses behavioral properties qualitatively similar to a non-linear oscillator. That constrained collectives of muscles exhibit a likeness to oscillatory mechanisms was intuited many years ago by Bernstein (1947) (see Greene, 1978) and advanced on empirical grounds by Fel'dman's mechanographic analysis of arm movements (Asatryan & Fel'dman, 1965; Fel'dman, 1966a, 1966b). More recent data (see Section 4) have borne out Fel'dman's work, in detail, and led to an elaboration of the potential advantages that such a view offers in terms of economy of neural control (cf. Greene, 1978; Kelso & Holt, 1980).

The present research is continuous with the above cited work, which suggests that muscles acting at a joint--to a first approximation--exhibit behavior qualitatively similar to a mass-spring system. But it also recognizes the universal fact that biological devices must necessarily incorporate non-linear features if local thermodynamic losses are to be compensated for and behavior sustained. The mathematical description for persistent cyclical operation is known as the limit cycle, and the present experiments have shown that patterns of coordination between the limbs (and within a single limb) can be accurately predicted from limit cycle properties. To summarize briefly, we have shown a clear-cut tendency for cycling limbs to maintain fixed amplitude and frequency under a variety of experimental conditions (brief and constantly applied load perturbations, the presence and absence of fixed mechanical constraints, different external driving frequencies). In addition, the tight phasic and timing relationships observed before and after imposed perturbations, as well as the demonstration that limbs cycling at different frequencies reveal non-arbitrary subharmonic relationships, attests strongly to the entrainment property of non-linear, limit cycle oscillators.

Though motivated by quite different concerns, the present experimental results fit rather well with Fel'dman's most recent work (cf. Fel'dman, 1980), in which subjects performed rhythmic movements of the right elbow joint at different frequencies and under various loading conditions. Like us, Fel'dman found that after unloading, rhythmic movement was preserved without visible change of either phase or frequency (see Fel'dman, 1980, Figure 5). According to Fel'dman, this result indicates that the "central generator" is independent of afferent influences created by the unloading. The latter conclusion is compatible with new data on phasic movements of the ankle joint (cf. Gottlieb & Agarwal, in press). When torques were applied at various points during movement initiation and execution, both early, myotatic (40-100 msec) and later, post-myotatic (100-300 msec) reflex responses to load changes were suppressed. Such was the case regardless of whether the phasic movements were of a slow, ramp nature or a fast, ballistic nature. Gottlieb and Agarwal (in press) suggest that there is substantial "preprogramming" of both types of movement, and that phasic movements in general are not assisted by effective load-compensating reflex mechanisms.

While the neurophysiological findings of Fel'dman and Gottlieb and Agarwal provide encouraging support for the present work, their focus is on the details of the neuromuscular response to externally imposed loads on a single joint. In contrast, the present experiments used load perturbations (among other manipulations) as a tool to discover patterns of coupling between the limbs, predictable from the properties of non-linear oscillators (see also Yamanishi, Kawato, & Suzuki, 1979, 1980, who use discrete visual, verbal, or manual events to perturb cyclic finger tapping). Of the latter, we have argued that the property of entrainment may be most significant for a viable theory of coordination.9

As an example, entrainment (often under a different name, see Section 6), may well account for the stable phase and timing relationships that characterize the various gaits of animal locomotion (for an excellent review, see Gallistel, 1980). It now appears that interneurons--at least in the central nervous systems of crayfish and cockroach--carry coupling or entraining signals that enable one limb to coordinate with another (cf. Delcomyn, 1980; Stein, 1976). The role of such "coordinating neurons" is currently being explored within a conceptual framework provided by the mathematical theory of coupled oscillators (Stein, 1976, 1977).

While we are sympathetic to the foregoing enterprise, we are also reminded of Davis' (1976) warning that properties of command, oscillation, coordination, and so on are not invested in any specific neuron. Coordination and oscillation are functions that reflect the interaction of cells, and are most correctly thought of as emergent properties. Davis (1976) provides several examples in which a particular function arises from a neuronal network even though no single neuron within the network possesses that function.

The theme that functions like coordination are emergent, a posteriori consequences of systematic interactions among cells (or muscles), as opposed to a priori prescriptions invested in a single cell (or program), is consonant with the dynamical perspective offered here (see also Fentress, 1976, for a similar "relational dynamics" perspective). For the physical theory of living systems (homeokinetics), entrainment is the chief mode of cooperation among self-sustaining oscillators; it is an emergent, self-organizing process in the sense that a collection of mutually entrained oscillators functions as a single unit. Therein lies its appeal, of course, for a principled solution to the degrees of freedom problem.

A further, and by now self-evident consequence of the homeokinetic view concerns the role of oscillation. It has long been recognized that cyclicity lies at the heart of biological functioning (cf. Goodwin, 1970, p. 8). Yet in the domain of movement, it has been commonplace (with certain notable exceptions) to consider fluctuating events as mere nuisances--as unwanted sources of variability. It has been easier to model control in terms of quasilinear servosystems than to search for ways in which oscillation may be exploited. In closed-loop servomechanisms, oscillation is undesirable because it means that there is a discrepancy between the input and the reference level or set point, and hence the system is unstable. However, a more important role for oscillatory processes is in non-linear systems that do not possess reference levels but that attain stability by virtue of entrainment.

The present studies and our preliminary work on the interaction of speech and gesture (Kelso, Tuller, & Harris, in press) are motivated by the latter theme. Relatedly, the approach to understanding motor control pursued in these experiments follows precisely the line of research proposed recently by Delcomyn (1980) following an extensive review of the neural basis for rhythmic behavior in lower phyla. Delcomyn (1980) identifies three problem areas, answers to which "...will bring neuroscientists much closer to the ultimate goal of understanding how nervous systems function" (p. 497). These are: "(i) the nature of an oscillator, (ii) the interaction of oscillators, (iii) the way in which sensory inputs interact with oscillators and their output to shape the final motor output" (p. 497). The arguments proposed herein suggest that principled solutions to the foregoing problems may lie in physical biology, particularly in Homeokinetics and Dissipative Structure theory (see Kugler et al., 1982; Prigogine, Note 4; Yates, in press, for comparisons and contrasts).

In our concluding remarks, we wish to make explicit one further contrast (that may already be apparent) between the present view and current concepts in the motor behavior area. As we noted at the beginning of this article, the conventional view attributes the regularities we observe in movement to an explicit, a priori prescription. But in an oscillator ensemble, there are no fixed dominance relationships in the sense that a program or reference level stands in a fixed, autocratic relation to the muscles responsible for implementation. There are different modes of interaction (e.g., frequency and amplitude modulation) and there may be "preferred" phase relationships under conditions of maximum coupling. A wide variety of behavioral patterns emerge from these interactions; there is structure and presumably a complex network of neuronal interconnections to support such cooperative phenomena, but strictly speaking, there is no dominance relation. If this view is correct, coordination is not prescribed by anything; it is more properly viewed as an emergent consequence of the dynamical behavior of a system whose design is fundamentally periodic.

REFERENCE NOTES

1. Kelso, J. A. S., Putnam, C. A., & Goodman, D. *Human interlimb coordination. 1. A kinematic analysis*. Manuscript in preparation, 1982.
2. Goodman, D., & Kelso, J. A. S. *Human interlimb coordination. 2. Temporal effects on initiation*. Manuscript in preparation, 1982.
3. Goodman, D., Rubin, P., & Kelso, J. A. S. *A generalized system for the capture and analysis of movement data*. Invited paper presented at the workshop on The Use of Minicomputers in Motor Behavior Research, North American Society for the Psychology of Sport and Physical Activity, Boulder, Colorado, 1980.
4. Prigogine, I. *Multiple time perspectives in dissipative structures*. Paper presented at the annual meeting of the American Psychological Association, Montreal, September, 1980.
5. Kelso, J. A. S. On the oscillatory basis of movement. Paper presented at Psychonomic Society, Philadelphia, November, 1981. Also in *Bulletin of the Psychonomic Society*, 1981, *18*, 63.
6. Schmidt, R. A. Personal communication, December, 1980.

REFERENCES

Adams, J. A. A closed-loop theory of motor learning. *Journal of Motor Behavior*, 1971, *3*, 111-150.

Adams, J. A. Feedback theory of how joint receptors regulate the timing and positioning of a limb. *Psychological Review*, 1977, *84*, 504-523.

Andronov, A., & Chaiken, C. E. *Theory of oscillations*. Princeton, N.J.: Princeton University Press, 1949.

Asatryan, D. G., & Fel'dman, A. G. Functional tuning of the nervous system with control of movement or maintenance of a steady posture - I. Mechanographic analysis on the work of the joint on execution of a postural task. *Biophysics*, 1965, *10*, 925-935.

Aschoff, J. Circadian rhythms: General features and endocrinological aspects. In D. Krieger (Ed.), *Endocrine rhythms*. New York: Raven Press, 1979.

Bellman, R. E. *Adaptive control processes: A guided tour*. Princeton, N.J.: Princeton University Press, 1961.

Berlinski, D. *On systems*. Boston: MIT Press, 1976.

Bernstein, N. A. *On the construction of movements*. Moscow: Medgiz, 1947. (Monograph: In Russian)

Bernstein, N. A. *The coordination and regulation of movements*. London: Pergamon Press, 1967.

Brazelton, T. B., Koslowski, B., & Main, M. The origin of reciprocity: The early mother-infant interaction. In M. Lewis & L. Rosenblum (Eds.), *The effect of the infant on its caregiver*. New York: John Wiley, 1974.

Brookhart, J. M., & Stein, P. S. G. *Locomotion*. Quincy, Mass.: Grass Instrument Company, 1980.

Cohen, L. Interaction between limbs during bimanual voluntary activity. *Brain*, 1970, *93*, 259-272.

Cooke, J. D. The organization of simple, skilled movements. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior*. Amsterdam: North Holland, 1980.

Davis, G. Organizational concepts in the central motor networks of invertebrates. In R. M. Herman, S. Grillner, P. S. G. Stein, & D. G. Stuart (Eds.), *Neural control of locomotion*. New York: Plenum, 1976.

Delcomyn, F. Neural basis of rhythmic behavior in animals. *Science*, 1980, *210*, 492-498.

Fel'dman, A. G. Functional tuning of the nervous system with control of movement or maintenance of a steady posture. II. Controllable parameters of the muscles. *Biophysics*, 1966, *11*, 565-578. (a)

Fel'dman, A. G. Functional tuning of the nervous system with control of movement or maintenance of a steady posture. III. Mechanographic analysis of execution by man of the simplest motor tasks. *Biophysics*, 1966, *11*, 766-775. (b)

Fel'dman, A. G. Superposition of motor programs--II. Rapid forearm flexion in man. *Neuroscience*, 1980, *5*, 91-95.

Fentress, J. C. Dynamic boundaries of patterned behavior: Interaction and self-organization. In P. P. G. Bateson & R. A. Hinde (Eds.), *Growing points in ethology*. Cambridge: Cambridge University Press, 1976.

Fowler, C. *Timing control in speech production*. Bloomington, Ind.: Indiana University Linguistics Club, 1977.

Fowler, C. A., Rubin, P., Remez, R. E., & Turvey, M. T. Implications for speech production of a general theory of action. In B. Butterworth (Ed.), *Language production*. New York: Academic Press, 1980.

Gallistel, C. R. *The organization of action: A new synthesis*. New York: Erlbaum, 1980.

Goodwin, B. Biological stability. In C. H. Waddington (Ed.), *Towards a theoretical biology*. Chicago: Aldine, 1970.

Gottlieb, G. L., & Agarwal, G. C. Response to sudden torques about the ankle in man: III. Suppression of stretch-evoked responses during phasic contraction. *Journal of Neurophysiology*, in press.

Graham, D. A model for the control of coordinated leg movements in free walking insects. *Biological Cybernetics*, 1977, *26*, 187-198.

Greene, P. H. Problems of organization of motor systems. In R. Rosen & F. Snell (Eds.), *Progress in theoretical biology* (Vol. 2). New York: Academic Press, 1972.

Greene, P. H. *Technical Report* (78-12). Illinois Institute of Technology: Department of Computer Science, 1978.

Hanson, F. E. Comparative studies of firefly pacemakers. *Federation Proceedings*, 1978, *37*, 2158-2164.

Hollerbach, J. J. *An oscillation theory of handwriting*. Cambridge, Mass.: MIT Artificial Intelligence Laboratory, 1980.

Holst, E. von. On the nature of order in the central nervous system (1937). In *The behavioral physiology of animals and man: The collected papers of Erich von Holst*. Coral Gables: University of Miami Press, 1973.

Holst, E. von. Relative coordination as a phenomenon and as a method of analysis of central nervous functions (1939). In *The behavioral physiology of animals and man: The collected papers of Erich von Holst*. Coral Gables: University of Miami Press, 1973.

Holst, E. von. *The behavioral physiology of animals and man*. Coral Gables: University of Miami Press, 1973.

Iberall, A. Periodic phenomena in organisms seen as non-linear systems. *Theoria to theory*, 1970, *4*, 40-53.

Iberall, A. S. A field and circuit thermodynamics for integrative physiology: I. Introduction to general notions. *American Journal of Physiology/Regulatory, Integrative, & Comparative Physiology*, 1977 *2*, R171-R180.

Iberall, A. S. A field and circuit thermodynamics for integrative physiology: III. Keeping the books - a general experimental method. *American Journal of Physiology/Regulatory, Integrative, & Comparative Physiology*, 1978, *3*, R85-R97.

Iberall, A. S., & McCulloch, W. S. The organizing principle of complex living systems. *Transactions of the American Society of Mechanical Engineers*, June, 1969, 290-294.

Iberall, A. S., Soodak, H., & Hassler, F. A field and circuit thermodynamics for integrative physiology: II. Power and communicational spectroscopy in biology. *American Journal of Physiology/Regulatory, Integrative, & Comparative Physiology*, 1978, *3*, R3-R19.

Katchalsky, A., & Curran, P. F. *Nonequilibrium thermodynamics*. Cambridge, Mass.: Harvard University Press, 1967.

Keele, S. W. A behavioral analysis of motor control. In V. Brooks (Ed.), *Handbook of physiology: Motor control*. Washington D.C.: American Psychological Society, 1981.

Kelso, J. A. S. Motor control mechanisms underlying human movement reproduction. *Journal of Experimental Psychology*, 1977, *3*, 529-543.

Kelso, J. A. S. Contrasting perspectives in order and regulation in movement. In J. Long & A. Baddeley (Eds.), *Attention and performance IX*. Hillsdale, N.J.: Erlbaum, 1981.

Kelso, J. A. S., & Holt, K. G. Exploring a vibratory systems analysis of human movement production. *Journal of Neurophysiology*, 1980, *43*, 1183-1196.

Kelso, J. A. S., Holt, K. G., & Flatt, A. E. The role of proprioception in the perception and control of human movement: Toward a theoretical reassessment. *Perception & Psychophysics*, 1980, *28*, 45-52.

Kelso, J. A. S., Holt, K. G., Kugler, P. N., & Turvey, M. T. On the concept of coordinative structures in dissipative structures: II. Empirical lines of convergence. In G. E. Stelmach (Ed.), *Tutorials in motor behavior*. New York: North-Holland Publishing Co., 1980.

Kelso, J. A. S., Southard, D. L., & Goodman, D. On the nature of human interlimb coordination. *Science*, 1979, *203*, 1029-1031. (a)

Kelso, J. A. S., Southard, D. L., & Goodman, D. On the coordination of two-handed movements. *Journal of Experimental Psychology: Human Perception and Performance*, 1979, *5*, 229-238. (b)

Kelso, J. A. S., Tuller, B. H., & Harris, K. S. A 'dynamic pattern' perspective on the control and coordination of movement. In P. MacNeilage (Ed.), *Motor control of speech production*. New York: Springer-Verlag, in press.

Klapp, S. T. Doing two things at once: The role of temporal compatibility. *Memory & Cognition*, 1979, *7*, 375-381.

Klima, E. S., & Bellugi, U. *The signs of language*. Cambridge, Mass.: Harvard University Press, 1979.

Kugler, P. N. K., Kelso, J. A. S., & Turvey, M. T. On the concept of coordinative structures as dissipative structures: I. Theoretical lines of convergence. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior*. New York: North-Holland Publishing Co., 1980, 1-47.

Kugler, P. N., Kelso, J. A. S., & Turvey, M. T. On coordination and control in naturally developing systems. In J. A. Scott Kelso & J. C. Clark (Eds.), The development of movement control and coordination. New York, London: John Wiley, 1982.

Laabs, G. J. Retention characteristics of different cues in motor short-term memory. Journal of Experimental Psychology, 1973, 100, 168-177.

Marteniuk, R. G., & MacKenzie, C. L. A preliminary theory of two-handed coordinated control. In G. E. Stelmach & J. Requin (Eds.), Tutorials in motor behavior. Amsterdam: North Holland, 1980.

Menaker, M. Physiological and biochemical aspects of circadian rhythms. Federation Proceedings, 1976, 35, 2325.

Minorsky, N. Nonlinear oscillations. Princeton, N.J.: Van Nostrand, 1962.

Morowitz, H. J. Energy flow in biology: Biological organization as a problem in thermal physics. New York: Academic Press, 1968.

Morowitz, H. J. Energy flow in biology. Woodbridge, Conn.: Oxbow Press, 1979.

Oatley, K., & Goodwin, B. W. The explanation and investigation of biological rhythms. In W. P. Colquhoun (Ed.), Biological rhythms and human performance. New York: Academic Press, 1971.

Pattee, H. H. Dynamic and linguistic modes of complex systems. International Journal of General Systems, 1977, 3, 259-266.

Pavlidis, T. Biological oscillators: Their mathematical analysis. New York: Academic Press, 1973.

Pew, R. W. Human perceptual motor-performance. In B. H. Kantowitz (Ed.), Human information processing: Tutorials in performance and cognition. New York: Lawrence Erlbaum Associates, 1974.

Polit, A., & Bizzi, E. Processes controlling arm movements in monkeys. Science, 1978, 201, 1235-1237.

Prigogine, I. Order through fluctuation: Self-organization and social systems. In E. Jantsch & C. H. Waddington (Eds.), Evolution and consciousness: Human systems in transition. Reading, Mass.: Addison-Wesley, 1976.

Prigogine, I., & Nicolis, G. Biological order, structure and instabilities. Quarterly Reviews of Biophysics, 1971, 4, 107-148.

Saunders, P. T. An introduction to catastrophe theory. Cambridge: Cambridge University Press, 1980.

Schepelmann, F. Rhythmic patterns of motor activity after lesions of the central nervous system in man. Acta Neurochirurgica, 1979, 69, 153-189.

Schmidt, R. A. A schema theory of discrete motor skill learning. Psychological Review, 1975, 82, 225-260.

Schmidt, R. A. On the theoretical status of time in motor program representations. In G. E. Stelmach & J. Requin (Eds.), Tutorials in motor behavior. Amsterdam: North Holland, 1980.

Schroedinger, E. What is life? London and New York: Cambridge University Press, 1945.

Searle, J. R. Minds, brains and programs. The Behavioral and Brain Sciences, 1980, 3, 417-457.

Shaffer, L. H. Analyzing piano performance: A study of concert pianists. In G. E. Stelmach & J. Requin (Eds.), Tutorials in motor behavior. Amsterdam: North Holland, 1980.

Shapiro, D. C., & Schmidt, R. A. Schema theory: Recent evidence and developmental implications. In J. A. S. Kelso & J. E. Clark (Eds.), The development of movement control and coordination. London: John Wiley, 1982.

Shik, M. L., & Orlovskii, G. N. Coordination of the limbs during running of the dog. *Biophysics*, 1965, *10*, 1148-1159.

Sollberger, A. *Biological rhythm research*. Amsterdam: Elsevier, 1965.

Soodak, H., & Iberall, A. S. Homeokinetics: A physical science for complex systems. *Science*, 1978, *201*, 579-582.

Stein, P. S. G. Mechanisms of interlimb phase control. In R. M. Herman, S. Grillner, P. S. G. Stein, & D. G. Stuart (Eds.), *Neural control of locomotion*. New York: Plenum, 1976.

Stein, P. S. G. Application of the mathematics of coupled oscillator systems to the analysis of the neural control of locomotion. *Federation Proceedings*, 1977, *36*, 2056-2059.

Stelmach, G. E. Retention of motor skills. *Exercise and sports sciences reviews* (Vol. 2). New York: Academic Press, 1974.

Thom, R. In D. H. Fowler (Trans.), *Structural stability and morphogenesis*. Reading, Mass.: Benjamin, Inc., 1975.

Turvey, M. T. Preliminaries to a theory of action with reference to vision. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting and knowing: Toward an ecological psychology*. Hillsdale, N.J.: Erlbaum, 1977.

Turvey, M. T. Clues from the organization of motor systems. In U. Bellugi & M. Studdert-Kennedy (Eds.), *Signed and spoken language: Biological constraints on linguistic form*. Basel: Verlag Chemie, 1980.

Turvey, M. T., Shaw, R. E., & Mace, W. Issues in the theory of action: Degrees of freedom, coordinative structures and coalitions. In J. Requin (Ed.), *Attention and performance VII*. Hillsdale, N.J.: Lawrence Erlbaum, 1978.

Wiener, N. Perspectives in cybernetics. In N. Wiener & J. P. Schade (Eds.), *Cybernetics of the nervous system*. Amsterdam: Elsevier, 1965.

Yamanishi, J., Kawato, M., & Suzuki, R. Studies on human finger tapping neural networks by phase transition curves. *Biological Cybernetics*, 1979, *33*, 199-208.

Yamanishi, J., Kawato, M., & Suzuki, R. Two coupled oscillators as a model for coordinated finger tapping of both hands. *Biological Cybernetics*, 1980, *37*, 219-225.

Yates, F. E. Physical biology: A basis for modeling living systems. *Journal of Cybernetics and Information Science*, in press.

Yates, F. E. Physical causality and brain theories. *American Journal of Physiology*, 1980, *238*, R277-R290.

Yates, F. E., & Iberall, A. S. Temporal and hierarchical organization in biosystems. In J. Urquart & F. E. Yates (Eds.), *Temporal aspects of therapeutics*. New York: Plenum, 1973, 17-34.

Yates, F. E., Marsh, D. J., & Iberall, A. S. Integration of the whole organism: A foundation for a theoretical biology. In J. A. Behnke (Ed.), *Challenging biological problems: Directions towards their solution*. New York: Oxford University Press, 1972, 110-132.

FOOTNOTES

[1] All current theories of regulation and control--mathematical systems theory, automata theory, cybernetics--though different in detail are alike in likening man to a machine; it is to this issue that we address limitations.

[2] Examples of biological escapements can be found in numerous metabolic processes such as the Krebs cycle where release of ATP is triggered by the metabolic demands of the cell.

[3] Strictly speaking, regulators and servomechanisms are different concepts even though it is common in the motor control area to hear them used synonymously. In the former, the reference value remains constant; in the latter, the reference value varies continually.

[4] Synchronization and entrainment are often used synonymously in the literature. Strictly speaking, synchronization is that state that occurs when both interval and phase of coupled oscillators are matched exactly; entrainment refers to the matching of intervals, but one oscillator may lead or lag the other.

[5] Interestingly, in American Sign Language (Klima & Bellugi, 1979) the constraints observed in the Kelso et al. experiments are omnipresent. According to Klima and Bellugi (1979): "If both hands move independently during a sign's articulation, then the two hands must exhibit identical hand configurations; the points of articulation are severely constrained with respect to one another (they must be in the same location or on the same horizontal or vertical plane); and the movements of the two hands must be the same (whether performed simultaneously or in alternation). The symmetry constraint thus specifies that in a two-handed sign, if both hands move and are active, they must perform roughly the same motor acts" (p. 64).

[6] On very few occasions, due to malfunctions in analog-to-digital conversion, a subject's data could not be registered for future analysis. As the Tables indicate however, we still had a sizable number of observations on which to base statistical comparisons.

[7] Two things should be emphasized in this section. The first is that all we offer here is a demonstration of a phenomenon that is coherent with the theoretical picture. A more formal analysis of sub-harmonic entrainment is being undertaken and some preliminary results have been presented (Kelso, Note 5). Second, as pointed out by R. A. Schmidt (Note 6) the pattern classified as 3:2 in Figure 5 is not 3:2 in the musical sense. In fact the "3" part is really a blank-move-move, which is the same as the "2" part. Again, this analysis attests to the difficulty subjects have in performing independent rhythms with the hands.

[8] In fact, more recent work (reported in Kelso, Note 5) indicates that if the two limbs operating out-of-phase (flexion in one and extension in the other) are driven at a certain critical frequency, they will change phase abruptly (in less than a cycle period) to an "in-phase" pattern. The 180 degree transition in phase brought about by a scalar increase in driving frequency is perhaps the principal feature of one of the simplest non-linear

systems modelled by the Duffing equation. Thus, through a continuous change in one variable (e.g., frequency), a discontinuous change in another (e.g., phase) can be observed. That is, an elementary cusp catastrophe or bifurcation occurs (Saunders, 1980; Thom, 1975). Importantly, there is no *a priori* prescription for the changed mode of organization anywhere.

[9] Norbert Wiener (1965), in a lecture prepared posthumously, discusses the benefits of entrainment as a non-linear, self-organizing phenomenon in biology and engineering. There are indications in this paper that Wiener was seeking solutions to nervous system organization in the dynamics of non-linear oscillators--with entrainment as a major feature. This view, as we have argued here, is rather different from the closed-loop, cybernetic theory that Wiener popularized.

MOTOR CONTROL: WHICH THEMES DO WE ORCHESTRATE?*

J. A. S. Kelso+ and E. L. Saltzman

When we are confronted with a living system, whose design is mysterious and whose optimizations are obscure, it is no easy task--as Professor Stein reminds us--to arrive at an answer to the question he has posed in this paper. Stein's target article is an important contribution for two main reasons. The first, which we shall mention only in passing, is that it is likely to provide much debate on what the controlled variables might be; moreover, it will force those who find this a burning issue to put their cards on the table. The second, and we feel more important reason, is that the paper poses a question--"What muscle variable(s) does the nervous system control...?"--whose very nature raises questions about the strategies neuroscience uses to investigate problems of control and coordination of movement. In our commentary we will focus on some of the (not so) implicit assumptions behind the question posed by Stein; if nothing else we hope to heighten sensitivity to some of the issues involved and (perhaps) to force the neuroscientist to consider his/her epistemology.

There are a couple of questionable assumptions in Stein's approach--at least as reflected in this article. The first is that control is the province of the nervous system; the second is that it is muscle variable(s) that are controlled. We shall examine each assumption in turn and their consequent ramifications for elucidating principles of coordination and control. In addition, we shall point to one notable omission in the author's list of candidates for control, and in our final remarks take up Stein's invitation to advance, albeit briefly, an alternative position to the control theoretic stance that he advocates here. Although it may, for the author, seem "...natural to assess performance...in similar terms to those applied to motors or other devices which produce movement," we assert that there are certain fundamental differences between living systems and machines (apart from structure) that render such a strategy not only dubious but highly unnatural. Most so-called "machine theories" regard biological control as a technical or engineering problem in which the many degrees of freedom to be regulated are a "curse" (Bellman, 1961). In contrast, there are contemporary physical theories yet to be explored fully in the domain of movement that

---

*An invited commentary on R. B. Stein "What muscle variable(s) does the nervous system control?" The Behavioral and Brain Sciences, in press.

+Also Departments of Psychology and Biobehavioral Sciences, University of Connecticut, Storrs.

Acknowledgment. Preparation of this paper and some of the research discussed herein were supported by NINCDS Grant NS-13617, BRS Grant RR-05596 and a PHS Post-Doctoral fellowship HD05676 to Elliot L. Saltzman.

[HASKINS LABORATORIES: Status Report on Speech Research SR-69 (1982)]

consider many degrees of freedom and non-linearities to be a requisite (not a source of "complication") for the emergence of coherent phenomena. The physical strategies being uncovered (principally in the form of Iberall and colleagues' Homeokinetic Theory, Prigogine and colleagues' Dissipative Structure Theory, Haken's Synergetics and Morowitz's Bioenergetics; see the final part of this commentary for references, and Kugler, Kelso, and Turvey, 1982, for contrasts among these theories and their application to control and coordination) stress autonomy, self-organization and evolution of function as system attributes--attributes that already appeal to some neuroscientists (cf. Katchalsky, Rowland, & Blumenthal, 1974; Llinas & Iberall, 1977; Szentagothai, 1978). Though we can classify ourselves as, at best, informed amateurs in this area, we believe these system attributes will prove difficult for the student of movement to ignore. A consideration of the assumptions behind the question posed by Stein may allow us to ground this claim more firmly.

Control as the Province of the Nervous System

It was surely one of Bernstein's (1967) most significant contributions (and he made many that have still to be appreciated) that control and coordination are not reducible to the orchestration of neural signals to and from the motor apparatus. Stein appears to recognize this fact in several places (e.g., in his discussions of stiffness, and his awareness of the possibility that energy fluxes may shape control), but the paper as a whole shows little appreciation of it. In fact, the predominant methodology in the studies cited by the author dictates that the organism and its parts are quiescent until mechanically or electrically stimulated. Obviously, we do not wish to be interpreted as saying that such a methodology has not proved useful in many cases or that the effects observed are not real. But control involves more than reactivity, and its analysis goes far beyond the deterministic input-output approach espoused in Stein's paper. One wonders to what extent the results of the studies cited by Stein, many of which involve single muscles, in non-intact preparations, can be generalized to normal movements in organisms continually interacting with their environments. There are a number of grounds for expressing skepticism on this issue (cf. Bernstein, 1967). To be blunt, an unequivocal relation between neural impulses to muscles and resulting movement does not, and cannot exist (see Benati, Gaglio, Morasso, Tagliasco, & Zaccaria, 1980; Boylls, 1975; Kelso & Holt, 1980; Saltzman, 1979; Turvey, Shaw, & Mace, 1978, for anatomical, mechanical, and physiological sources of functional non-univocality or indeterminacy).

Consider, for example, the task of maintaining the elbow at a steady state angle of 180 degrees (i.e., the elbow in full extension). The orientation of the arm in the gravity field determines not only the relative contributions of muscle and gravity torques at the desired elbow angle, but also the stability properties of this equilibrium configuration. When the arm is in a downward vertical orientation, the elbow angle is stable; if the elbow is perturbed (flexed), it will return to equilibrium due to the stable restoring torque of gravity. No muscle activity is required in this case. If, however, the arm is in an upward vertical orientation, the elbow angle is unstable since gravity plays a destabilizing role in this configuration. If the equilibrium angle is to be restored, muscle activity is required to provide the stabilizing restorative torque. Thus, the relative contributions

118

of gravity and muscle stiffnesses for a stable equilibrium angle vary with the arm's orientation in the gravity field. In short, the nervous system "controls" only to the extent that it complements the force field of the environment. Its role is better envisaged as exploitative rather than injunctive.

On the Selection of Analytic Units

In the target article, Stein has attempted to pinpoint the variables used by the nervous system to control muscular activity during the performance of sensorimotor tasks. Such tasks might involve a limb that moved, remained stationary, or exerted forces and torques at the limb-environment interface. It is important to recognize, however, that such controlled variables are defined only in the context of the organizational structure(s) comprising the system-to-be-controlled. These organizational structures are defined functionally at a higher level of description than that of single muscles or joints. Specifically, they are defined at the relatively abstract task level and, as mentioned in the section above, span the dynamic system composed of both organism and environment. In solving the problem posed by a task, the nervous system is the indispensable medium through which the requisite limb organization can emerge. The limbs (or any set of articulators) thus become different types of functionally defined, special purpose devices for different types of tasks.

Although the immediately preceding statements may seem trivial at first glance, they reveal a perspective that has decidedly non-trivial implications for how we approach the problem of controlled variables. More specifically, this perspective leads us to place significant constraints on our selection of analytic units of behavior. Professor Stein, for example, limits his analysis to "...simple physical variables appropriate to single muscles or groups of muscles acting normally around a joint." Few scientists would disagree that some decomposition of the system is necessary for analytic purposes. However, the unit of analysis should not be casually or arbitrarily chosen, at least if the ultimate goal is to understand control in animals (not simply in a single joint). Our point can be made through an example from physics (cf. Rosen, 1978).

It is well known that the three-body problem defies an analytic solution in closed form: Whether the earth-sun-moon system is truly stable is an open question. Although it is possible to decompose the system into one-body and two-body subsystems that are completely tractable analytically, such a strategy does not facilitate obtaining a solution to the three-body problem. The reason is that the physical decomposition itself destroys the original dynamics. In order to solve the three-body problem, a new set of analytic units must be discovered that are defined by new observables, such that the partitioning of the system does not annihilate the original dynamics. As Rosen (1978) remarks, this partitioning will seem strange to us because we are used to selecting so-called "simple" units that correspond to some physical fractionation of the system. The point is that when we reduce or decompose the system, the greatest care must be taken in selecting the proper unit of analysis. It is most likely that "simplicity" (a term with an exceedingly slippery definition) will be neither the only, nor the chief criterion involved.

Returning to the domain of movement, the identification of appropriate units of analysis has long been a thorny issue, going back at least to Sherrington's reference to the reflex as "...a simple, if not a probable fiction" (Sherrington, 1906). More recently, Greene (1971) in echoing Nicolai Bernstein has remarked that much of our confusion in studying problems of coordination has arisen "from our limited ability to recognize the significant informational units of movement." However, there are signs (although only considered in passing in the target article) that some consensus may be drawing near. There has been a growing appreciation that individual muscles (or muscle variables) are not the proper units of analysis for discussing coordinated movements; rather, such movements are partitioned more naturally into collective functional units defined over groups of muscles and joints, within which component elements vary relatedly and autonomously (e.g., Boylls, 1975; Fowler, 1977; Kelso, Southard, & Goodman, 1979; Lestienne, 1979; Nashner, 1977; Saltzman, 1979; Szentagothai & Arbib, 1974; Talbott, 1979; Turvey, 1977). The Soviet school (e.g., Bernstein, 1967; Gelfand & Tsetlin, 1971; Shik & Orlovskii, 1976) refers to such muscle-joint groupings as linkages or synergies. These terms reflect an appreciation of the fundamental problem of control and coordination, namely that of constraining a complex system with many degrees of freedom to behave in a regular and orderly manner.

Synergies (or coordinative structures, cf. Easton, 1972; Kelso et al., 1979; Turvey et al., 1978) by definition are functionally specific units defined over groups of muscles and joints, which constrain the component elements to act together in a manner appropriate to the task at hand. Such muscle collectives are thought to share a common efferent and afferent organization and are deployable as relatively autonomous units in sensorimotor tasks (e.g., Boylls, 1975; Gelfand, Gurfinkel, Tsetlin, & Shik, 1971). Coordinative structures as functional units of control are currently undergoing rigorous analysis in a number of laboratories; they have been identified in various tasks and at different levels of analysis (cf. Kelso, 1981; Kelso, Tuller, & Harris, in press; Kugler, Kelso, & Turvey, 1980, for recent examples). Their chief feature rests in a mutable partitioning of component variables into those that preserve the structural ("topological") organization of movement (e.g., the relative timing and relative force properties of muscular events) and those that are capable of effecting scalar transformations on these qualitative structures. A theoretical rationale for coordinative structures has been offered (cf. Kelso, 1981; Kugler et al., 1980, 1982), focusing on those properties that distinguish movement patterns that exhibit structural stability from those that do not.

There are indications in the target article that Professor Stein prefers to sidestep the issue of functionally specific units of movement as not germane to his interests, and as one that pertains only to "multijoint movements" or the "large behavioral literature on complex patterned movement." However, he does not hesitate to negate arguments for length and stiffness control on the basis of "complex patterned movements" like speech or piano playing in the case of length, and walking in the case of stiffness. We welcome the functional argument in each case, although we note that for Stein it involves jumping rather precariously between muscle-joint levels (e.g., stiffness and length) and task levels of analysis (such as skiing and needle-threading). Though aware of the problem, Stein seems to apply a single muscle-joint unit of analysis generally to all types of complex multijoint tasks. Such an approach is at the same time too powerful and too arbitrary. It is too powerful because it allows descriptions of movement control that fail to distinguish between those acts that do occur and those acts that are physically possible but never occur. It is too arbitrary because single joint actions will rarely relate unequivocally to particular task functions. In short, when we deal with coordinated activity, we are dealing with task specific functional units whose degrees of freedom are constrained according to task demands, or more generally, to the mutual relationship between organism and environment.

Contrasting Views on the Origins of Order

Whenever we observe a regular and orderly phenomenon, it is always a temptation to assign responsibility to some device that is antecedent to, and causally responsible for, the said phenomenon. The device has available to it "representations" that have characteristics very much like the phenomenon we are trying to understand. As philosophers have often told us, "representations" require users with goals and interests (much like the animal itself) and so, when we assume their presence, we take out a loan on intelligence that must ultimately be paid back (cf. Dennett, 1978; Searle, 1980). We can bury our heads in the sand on this issue or we can approach the problem in a different way: one that asks not how control can be explained according to some *a priori* prescription *for* the system (such as the central representations and the cybernetic, negative feedback paradigms favored by Stein), but rather how control arises as an *a posteriori* consequence of the system's dynamical organization.

For example, imagine adopting the former, *prescriptive* strategy to a coherent biological phenomenon such as the schooling of fish. What we observe are individual fish behaving collectively in a highly coordinated manner. The "system" in this case has many degrees of freedom and exhibits an organized, seemingly wholistic structure. Adopting a prescriptive strategy, we might search the system for a "reference value" or a "central representation" that regulates the individual fish or the collective of fish, but it would make little sense to do so. These would be *special mechanisms* introduced by the unknowing observer to account for a poorly understood phenomenon. In fact, the highly coherent behavior of fish schooling can be accounted for with a fairly small set of key variables, such as "density" defined through the metric of fish length. When the average distance between nearest neighbors is less than one fish length (note that the metric is "intrinsic" and system-scaled; cf. Warren & Shaw, 1981), spacing between fish is schooled not random (cf. Okubo, 1980, for an in-depth analysis).

Although the details of collective fish behavior may seem far removed from the issues raised by Stein, there is, we think, an important message for the neuroscientist or psychologist. It is that an understanding of a complex, organizational phenomenon such as fish schooling rests with articulating the necessary and sufficient conditions for that organization to occur. More generally this approach entails a strategy that rejects the introduction of special mechanisms--as sources of explanation--before dynamics has been fully explored. Put another way, what can we, as students of movement, explain "for free" before we burden the nervous system with the onus of control?

In this regard, it is puzzling to us that Stein chooses to ignore a model whose dynamics obviate (or at least significantly reduce) the requirement for ongoing, computational control. If recent work is a guide, much may be gained through the identification of functional units of movement with nonlinear mass-spring systems. Although the model has received an uneven interpretation, its import for us is that it allows one to see the qualitative similarities between certain aspects of movement control (such as the ability to reach the same desired spatial location, with different trajectories and from variable initial conditions) and the behavior of a mass-spring system. Following our arguments expressed above, the beauty of the mass-spring model lies not in the literal parallel between a single muscle and a spring, but in the recognition that particular behaviors share--to a first approximation--the same abstract functional organization as a mass-spring system.

The intuition that a muscle-joint system is dynamically similar to a mass-spring system with controllable equilibrium length is due to Fel'dman (cf. Fel'dman, 1966, p. 771), and has undergone appropriate extension by a number of authors (e.g., Bizzi, Polit, & Morasso, 1976; Kelso, 1977; Kelso & Holt, 1980; Polit & Bizzi, 1978; Schmidt & McGown, 1980). The basic idea is that a given joint angle may be specified according to a set of muscle equilibrium lengths. Once these are specified, the joint will achieve and maintain a desired final angle at which the torques generated by the muscles sum to zero. Such a system exhibits the property of equifinality in that desired positions may be reached from various initial angles, and in spite of unforeseen perturbations encountered during the motion trajectory (see Kelso, Holt, Kugler, & Turvey, 1980 for review; but also Saltzman, 1979, for some cautionary notes). Fel'dman (1966, 1980) has further noted that stiffness at a joint may be specified in terms of agonist and antagonist equilibrium lengths even if the stiffness of these muscles is not itself controllable. In the Fel'dman model, joint stiffness covaries with the degree of agonist-antagonist co-contraction.

Two points for Stein emerge from this discussion. One concerns a sin of omission in that he includes the spring property of stiffness as a possible control variable, but neglects the related variable of equilibrium length. The other, perhaps more important issue warrants a little further development, because of its theoretical consequences. It is that in likening (to a first approximation) a constrained collective of muscles to a mass-spring system, the need to introduce externally imposed measurement, comparison, and control operations is reduced. Though we could describe a dynamical system like a mass-spring in terms of externally imposed reference levels and though we could mathematize it into canonical feedback form, little would be gained by doing so (cf. Yates, in press). A muscle collective qua spring system is intrinsically self-equilibrating: conserved values such as the equilibrium point emerge from the system's parameterization. More emphatically, in mass-spring systems (like schools of fish and functional groupings of muscles?) there is no need to introduce a "representation" anywhere.

Toward an Alternative Control Scheme

In our final comments we take up--in rather condensed fashion because of space limitations--Professor Stein's invitation to his critics to offer an alternative scheme to the one that he has put forward so authoritatively. We refer to an emerging theoretical view of movement control and coordination that has been expressed in two recent papers (Kugler et al., 1980, 1982) and that has also undergone some, as yet limited, empirical scrutiny (Kelso, Holt, Kugler, & Turvey, 1980; Kelso, Holt, Rubin, & Kugler, in press). Its origins (and to a large extent its appeal) lie in a unified treatment of cooperative phenomena at all scales of magnitude (cf. Haken, 1977). Hence it speaks to the important issue raised by Stein at the beginning of his article, namely that of relating levels of analysis. Moreover, the perspective is consonant with some of the themes introduced above and also may interface with evolving oscillator theoretic views of neural control not considered by Stein in the target article (e.g., Delcomyn, 1980).

A chief distinguishing feature of the view expressed here lies in the recognition that first and foremost, living systems belong to a class of physical systems that are open to fluxes of energy and matter (in contrast, cybernetic systems are closed to energy and matter exchange with their surrounds). The principal theories addressing such systems are Iberall's Homeokinetic Theory (e.g., Iberall, 1977, 1978; Soodak & Iberall, 1978; Yates & Iberall, 1973) and Prigogine's Dissipative Structure Theory (e.g., Nicolis & Prigogine, 1977; Prigogine, 1980). The former, in particular, addresses systemic phenomena in biology and elaborates, among other things, the conditions for persistence of function, autonomy, and self-organization. It represents a concerted effort to apply irreversible thermodynamics to living systems. A fundamental tenet is that in steady-state systems the flow of energy through the system plays an organizing role and that, following Morowitz's theorem, energy flow from a potential source to a lower order sink will lead to at least one cycle in the system (cf. Morowitz, 1968). Homeokinetic theory builds on the Barnard-Cannon principle of homeostasis (which contained no mechanism for preservation of conserved states); in the homeokinetic view, control is dynamically effected by means of coupled ensembles of limit cycle oscillatory processes. Limit cycles represent the only temporal stability for non-conservative, nonlinear systems; they resemble "squirt" systems that, by virtue of their design, are capable of making up for dissipative losses that occur in the drift toward equilibrium (see Yates & Iberall, 1973). The system's conserved values or equilibrium operating points are thought to be specified in the loose coupling of limit cycle processes. Limit cycles are manifestations of thermodynamic engines and quantize action (formally, the product of energy and time; cf. Iberall, 1978) at every level in the system.

As functional units of movement, ensembles of nonlinear limit cycle oscillators offer a number of attractive features for a principled account of coordination and control. Among these are their self-sustaining properties, their ability to operate independently of initial conditions, their stability in the face of moderate perturbations, and, perhaps most important for the theorist of movement, the properties of mutual entrainment and synchronization (cf. Minorsky, 1962; Winfree, 1980).

With respect to the issues raised by Stein, it is worth emphasizing that limit cycles are not special mechanisms *per se*. To observe spectrally distributed limit cycle regimes and for new spatiotemporal organizations to emerge, certain necessary conditions must exist. Among these are the presence of many interacting degrees of freedom, nonlinearities, a relatively constant source of potential energy and the requirement that energy be dissipated. Given such conditions and subject to critical scaling influences, constraints emerge that are capable of marshalling the free variables into coherent functions. Quadruped gait may be an example: When one stable movement pattern is driven beyond a critical value on a system-sensitive parameter, a bifurcation occurs and a new spatiotemporal pattern--a new stability--arises. In such a view, no explicit "gait selection" process is required (e.g., Gallistel, 1980).

To reiterate our main point, however, in the perspective offered here, order (control and coordination) is functionally specified in the system's dynamics. The radical claim, as Gibson (1979) once remarked, is that behavior is regular but there are no regulators. A less radical statement would be an affirmative answer to Yates' (1980) question to the readers of the *American Journal of Physiology*: Do [you] know of a serious effort to discharge the homunculus?

The spirit of the foregoing discussion leads us to raise one final issue. It is the growing intuition--stemming from theoretical considerations raised here and elsewhere (cf. Anderson, 1972)--that the problem of order in natural systems might be attacked more effectively by seeking out a single set of physical principles that can apply at all levels, rather than by positing different units of analysis at each level. One assumes nature operates with ancient themes. In this commentary we have tried to provide a flavor for the ones that neuroscience in general, and the field of motor control in particular, might consider worth orchestrating.

REFERENCES

Anderson, P. W. More is different. *Science*, 1972, *177*, 393-396.

Bellman, R. E. *Adaptive control processes: A guided tour*. Princeton, N.J.: Princeton University Press, 1961.

Benati, M., Gaglio, S., Morasso, P., Tagliasco, V., & Zaccaria, R. Anthropomorphic robotics. II. Analysis of manipulator dynamics and the output motor impedance. *Biological Cybernetics*, 1980, *38*, 141-150.

Bernstein, N. A. *The coordination and regulation of movements*. London: Pergamon Press, 1967.

Bizzi, E., Polit, A., & Morasso, P. Mechanisms underlying achievement of final head position. *Journal of Neurophysiology*, 1979, *39*, 435-444.

Boylls, C. C. A theory of cerebellar function with applications to locomotion. II. The relation of anterior lobe climbing fiber function to locomotor behavior in the cat. *COINS Technical Report 76-1*, Department of Computer and Information Science, University of Massachusetts, 1975.

Delcomyn, F. Neural basis of rhythmic behavior in animals. *Science*, 1980, *210*, 492-498.

Dennett, D. C. *Brainstorms: Philosophical essays on mind and psychology*. Montgomery, VT: Bradford Books, 1978.

Easton, T. A. On the normal use of reflexes. *American Scientist*, 1972, *60*, 591-599.

Fel'dman, A. G. Functional tuning of the nervous system with control of movement or maintenance of a steady posture. III. Mechanographic analysis of execution by man of the simplest motor tasks. *Biophysics*, 1966, 11, 766-775.

Fel'dman, A. G. Superposition of motor programs--II. Rapid forearm flexion in man. Neuroscience, 1980, 5, 91-95.

Fowler, C. Timing control in speech production. Bloomington, Ind.: Indiana University Linguistics Club, 1977.

Gallistel, C. R. The organization of action: A new synthesis. New York: Erlbaum, 1980.

Gelfand, I. M., Gurfinkel, V. S., Tsetlin, M. L., & Shik, M. L. Some problems in the analysis of movements. In I. M. Gelfand, V. S. Gurfinkel, S. V. Fomin, & M. L. Tsetlin (Eds.), Models of the structural-functional organization of certain biological systems. Cambridge, Mass.: MIT Press, 1971.

Gelfand, I. M., & Tsetlin, M. L. Mathematical modeling of mechanisms of the central nervous system. In I. M. Gelfand, V. S. Gurfinkel, S. V. Fomin, & M. L. Tsetlin (Eds.), Models of the structural-functional organization of certain biological systems. Cambridge, Mass.: MIT Press, 1971.

Gibson, J. J. The ecological approach to visual perceptions. Boston: Houghton-Mifflin, 1979.

Greene, P. H. Introduction. In I. M. Gelfand, V. S. Gurfinkel, S. V. Fomin, & M. L. Tsetlin (Eds.), Models of the structural-functional organization of certain biological systems. Cambridge, Mass.: MIT Press, 1971.

Haken, H. Synergetics: An introduction. New York: Springer-Verlag, 1977.

Iberall, A. S. A field and circuit thermodynamics for integrative physiology: I. Introduction to general notions. American Journal of Physiology/Regulatory, Integrative, & Comparative Physiology, 1977, 2, R171-R180.

Iberall, A. S. A field and circuit thermodynamics for integrative physiology: III. Keeping the books - a general experimental method. American Journal of Physiology/Regulatory, Integrative, & Comparative Physiology, 1978, 3, R85-R97.

Katchalsky, A. K., Rowland, V., & Blumenthal, R. Dynamic patterns of brain cell assemblies. Neurosciences Research Program Bulletin, 1974, 12(1).

Kelso, J. A. S. Motor control mechanisms underlying human movement reproduction. Journal of Experimental Psychology: Human Perception and Performance, 1977, 3, 529-543.

Kelso, J. A. S. Contrasting perspectives on order and regulation in movement. In J. Long & A. Baddeley (Eds.), Attention and performance IX. Hillsdale, N.J.: Erlbaum, 1981.

Kelso, J. A. S., & Holt, K. G. Exploring a vibratory systems analysis of human movement production. Journal of Neurophysiology, 1980, 43, 1183-1196.

Kelso, J. A. S., Holt, K. G., Kugler, P. N., & Turvey, M. T. On the concept of coordinative structures as dissipative structures: II. Empirical lines of convergence. In G. E. Stelmach & J. Requin (Eds.), Tutorials in motor behavior. New York: North-Holland Publishing Co., 1980.

Kelso, J. A. S., Holt, K. G., Rubin, P., & Kugler, P. N. Patterns of human interlimb coordination emerge from the properties of non-linear limit cycle oscillatory processes: Theory and data. Journal of Motor Behavior, in press.

Kelso, J. A. S., Southard, D. L., & Goodman, D. On the nature of human interlimb coordination. Science, 1979, 203, 1029-1031.

Kelso, J. A. S., Tuller, B. H., & Harris, K. S. A 'dynamic pattern' perspective on the control and coordination of movement. In P. MacNeilage

(Ed.), *Motor control of speech production*. New York: Springer-Verlag, in press.

Kugler, P. N., Kelso, J. A. S., & Turvey, M. T. On the concept of coordinative structures as dissipative structures: I. Theoretical lines of convergence. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior*. New York: North-Holland Publishing Co., 1980.

Kugler, P. N., Kelso, J. A. S., & Turvey, M. T. On coordination and control in naturally developing systems. In J. A. Scott Kelso & J. C. Clark (Eds.), *The development of human movement control and coordination*. New York: John Wiley, 1982.

Lestienne, F. Effects of inertial load and velocity on the braking process of voluntary limb movements. *Experimental Brain Research*, 1979, *35*, 407-418.

Llinas, R. R., & Iberall, A. A global model of neuronal command-control systems. *Biosystems*, 1977, *8*, 233-235.

Minorsky, N. *Nonlinear oscillations*. Princeton, N.J.: Van Nostrand, 1962.

Morowitz, H. J. *Energy flow in biology: Biological organization as a problem in thermal physics*. New York: Academic Press, 1968.

Nashner, L. M. Fixed patterns of rapid postural responses among leg muscles during stance. *Experimental Brain Research*, 1977, *30*, 13-24.

Nicolis, G., & Prigogine, I. *Self-organization in nonequilibrium systems: From dissipative structures to order through fluctuations*. New York: Wiley-Interscience, 1977.

Okubo, A. *Diffusion and ecological problems: Mathematic models*. Berlin: Springer-Verlag, 1980.

Polit, A., & Bizzi, E. Processes controlling arm movements in monkeys. *Science*, 1978, *201*, 1235-1237.

Prigogine, I. *From being to becoming: Time and complexity in the physical sciences*. San Francisco: W. H. Freeman & Co., 1980.

Rosen, R. *Fundamentals of measurement and representation of natural systems*. New York: North-Holland, 1978.

Saltzman, E. Levels of sensorimotor representation. *Journal of Mathematical Psychology*, 1979, *20*, 91-163.

Schmidt, R. A., & McGown, C. Terminal accuracy of unexpectedly loaded rapid movements: Evidence for a mass-spring mechanism in programming. *Journal of Motor Behavior*, 1980, *12*, 149-161.

Searle, J. Minds, brains and programs. *The Behavioral and Brain Sciences*, 1980, *3*, 417-457.

Sherrington, C. S. *The integrative action of the nervous system*. London: Constable, 1906.

Shik, M. L., & Orlovskii, G. N. Neurophysiology of locomotor automatism. *Physiological Reviews*, 1976, *56*, 465-501.

Soodak, H., & Iberall, A. S. Homeokinetics: A physical science for complex systems. *Science*, 1978, *201*, 579-582.

Szentagothai, J. A false alternative: Commentary on Pucetti and Dykes. *The Behavioral and Brain Sciences*, 1978, *3*, 367-368.

Szentagothai, J., & Arbib, M. A. (Eds.). Conceptual models of neural organization. *Neurosciences Research Program Bulletin*, 1974, *12*(3).

Talbott, R. E. Ferrier, the synergy concept, and the study of posture and movement. In R. E. Talbott & D. R. Humphrey (Eds.), *Posture and movement*. New York: Raven Press, 1979.

Turvey, M. T. Preliminaries to a theory of action with reference to vision. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting and knowing: Toward* an *ecological psychology*. Hillsdale, N.J.: Erlbaum, 1977.

Turvey, M. T., Shaw, R. E., & Mace, W. Issues in the theory of action: Degrees of freedom, coordinative structures and coalitions. In J. Requin (Ed.), *Attention and performance VII*. Hillsdale, N.J.: Lawrence Erlbaum, 1978.

Warren, W. H., & Shaw, R. E. Psychophysics and ecometrics. *The Behavioral and Brain Sciences*, 1981, *4*, 209-210.

Winfree, A. T. *The geometry of biological time*. New York: Springer-Verlag, 1980.

Yates, F. E. Getting the homunculus out of the head. *American Journal of Physiology*, 1980, *239*, R363-364.

Yates, F. E. Systems analysis of hormone action: Principles and strategies. In R. F. Goldberger (Ed.), *Biological regulation and development, Vol. III - Hormone action*. New York: Plenum, in press.

Yates, F. E., & Iberall, A. S. Temporal and hierarchical organization in biosystems. In J. Urquhart & F. E. Yates (Eds.), *Temporal aspects of therapeutics*. New York: Plenum, 1973.

EXPLORING THE NATURE OF MOTOR CONTROL IN DOWN'S SYNDROME*

Walter E. Davis+ and J. A. Scott Kelso++

Abstract. Following Asatryan and Fel'dman (1965), two experiments were conducted to describe the so-called invariant mechanical properties underlying movement control in Down's syndrome and normal subjects. The invariant characteristic is a curve on a graph of joint torque versus joint angle. The task required subjects to maintain a steady joint angle against an external load (torque). Torque was systematically changed via partial unloading in order to obtain torque by length (joint angle) functions at three separate initial joint angles. Instructions required subjects "not to intervene" when unloading occurred in Experiment 1 and to "tense" their muscles prior to unloading in Experiment 2. Both normal and Down's syndrome groups revealed systematic torque by length functions that might be expected according to a simple mass-spring system model. Although the gross organization of movement in Down's syndrome subjects was nearly the same as normals, important differences between the two groups were found. Down's syndrome subjects revealed underdamped motions relative to normals (as shown by differences in the degree of oscillation about the final equilibrium position) and were less able to regulate stiffness (as shown by differences in slope of the torque by angle functions in Experiment 2). We promote the notion that damping and stiffness may be sensitive indices of hypotonia--the most common description of neuromuscular deficiency in Down's syndrome.

There is very little research on the control and coordination of movement in Down's syndrome subjects. The few studies that do exist reveal that this population exhibits abnormal gait (James, 1974), slow movement responses (Berkson, 1960; Lange, 1970) and is less accurate on certain motor tasks (Frith & Frith, 1974) than their normal counterparts. However, questions concerning the underlying organization of motor control in Down's syndrome subjects remain unanswered.

---

*To appear in Journal of Motor Behavior, in press.

+The Motor Development Center, Kent State University, Kent, Ohio.

++Also Departments of Biobehavioral Sciences and Psychology, University of Connecticut, Storrs.

Acknowledgments. Portions of this paper represent a doctoral dissertation submitted to the University of Connecticut by Walter E. Davis, who is presently at The Motor Development Center, Kent State University, Kent, Ohio 44242. The work was supported by NIH grants AM 25814 and NS-13617, and BRS Grant RR-05596 to Haskins Laboratories. We thank Michael Turvey for sharing his insights with us on this problem and also for reviewing an earlier version. We are grateful to Philip Rubin who developed the "action" computer software for analyzing some of the data reported here.

The most often implicated motor deficiency in Down's syndrome subjects is hypotonia, a decrease or absence of muscle tone. Indeed, nearly all Down's syndrome infants are born with hypotonia, a condition that may contribute greatly to their delay in reaching motor milestones (cf. Cowie, 1970). However, no definitive statement can be made about hypotonia in Down's syndrome individuals beyond infancy (cf. Knight, Atkinson, & Hyman, 1967; Owens, Dawson, & Losin, 1971). To complicate things, the exact relationship between hypotonia and motor control is, as yet, unclear (cf. Neilson & Lance, 1978).

One model--for which there is accumulating empirical evidence--and which may allow us to understand movement organization in Down's syndrome, draws the analogy between the behavior of muscle collectives and a vibratory system (cf. Turvey, 1977). For example, Asatryan and Fel'dman (1965), in a now classic experiment had subjects establish steady-state joint angles against an external load. On each occasion, when the load was systematically reduced and the subjects instructed not to intervene, the joint angle changed. Resulting torque by length (joint angle) functions generated from different initial positions were parallel and non-intersecting as would be expected from a mass-spring model.

Additional evidence supporting a vibratory system model comes from research demonstrating equifinality in the muscle-joint systems of human subjects (e.g., Bizzi, 1980; Fel'dman, 1966b; Kelso, 1977; Kelso & Holt, 1980; Schmidt & McGowan, 1980, for reviews) and trained monkeys (Bizzi, Dev, Morasso, & Polit, 1978; Bizzi, Polit, & Morasso, 1976). Equifinality, as defined, refers to a system's ability to equilibrate based solely on its parameters and independent of initial conditions (von Bertalanffy, 1973).[1]

Following Asatryan and Fel'dman's (1965) analysis, the present study sought to describe the static mechanical properties of the muscle joint system of Down's syndrome subjects using a simple mass-spring equation: $F = -K (l - l_0)$, where F is an external force, -K is stiffness, l is the current length of the spring and l_0 is the length of the spring when no forces are acting on it. Experiment 1 was designed to determine the so-called invariant characteristics (cf. Fel'dman, 1980a) of the muscle-joint system of Down's syndrome subjects. This was achieved by examining the relationship between change in joint angle under conditions of partial unloading when subjects were instructed not to intervene voluntarily against the unloading, thereby holding the parameters of stiffness and zero length relatively constant. Under these conditions, any change in joint angle should be systematic with the change in the load (torque).

A second experiment examined the extent to which Down's syndrome subjects were capable of regulating stiffness by requiring subjects to tense their muscles voluntarily at particular joint angles prior to unloading. Asatryan and Fel'dman (1965) demonstrated that normal subjects could voluntarily increase the stiffness of the muscle-joint system as reflected by increases in the slope of the torque by length function. It is not known whether stiffness can be regulated voluntarily by Down's syndrome subjects. Qualitatively speaking, the Down's syndrome population is characterized by hypotonia (flaccidity), which may relate to muscle stiffness or to damping. Damping is the internal frictional force present in the system and is indicated by the extent of overshoot or oscillation about the equilibrium point. As this study shows, both stiffness and damping appear to be sensitive indices of the impaired performance of Down's syndrome subjects.

EXPERIMENT 1

Method

Subjects

The subjects in this study were seven Down's syndrome male students between 14 and 21 years of age who attended Celentano School for the developmentally handicapped in New Haven, Connecticut. These subjects were selected from among 30 subjects who participated in a previous study that employed limb localization movements (Davis & Kelso, Note 1). The subjects were selected on the basis of their ability to complete the task. The Celentano School administration reported an IQ range of 25-60 for the subjects involved in this study. A control group consisted of six normal male subjects selected from the subject pool listed at Haskins Laboratories in New Haven who were paid for their services.

Apparatus

The apparatus consisted of a 593c finger positioning device along with an associated electronics control package (see Kelso & Holt, 1980, for a description). The main parts of the apparatus were two movable arms each attached to a separate metal shaft mounted vertically on the top of an open box frame. Only the right hand was used in these experiments. The frame was mounted on a table 78.5 cm high.

The movements allowed by the positioning device were flexion and extension of the index finger about the metacarpophalangeal joint. The distal end of the moving finger was fitted with a plastic collar that slipped into an open-ended cylindrical support attached to the movable arm. The movable arm consisted of two parallel bars fitted perpendicularly into the metal shaft. A pointer was attached to the end of the movable arm and moved along a protractor calibrated in degrees. The apparatus was also equipped with padded adjustable braces with which to secure the subjects' wrist, hand, and remaining fingers and thumb during the movements.

The gear arrangement driven by a torque motor provided resistance (load) to the movable arm (and hence to the finger when placed in the cylinder attached to the arm) when current was supplied to the motor. The electronics control box allowed for regulation of the current supply and could be set in either of two modes. While in the first mode, which shall be referred to as "servo-torque control," a resistance (torque) acted on the finger whenever the finger deviated from the set servo-position. When the finger was in the servo-position, no resistance acted upon it. When in the second mode, referred to as "torque control," a constant resistance could be applied to the finger (settable in either direction). The amount of resistance (torque) was adjustable and could be set anywhere from 0 to 100% of the maximum torque available from the motor (81.6 ounce-inches).

Procedure

Procedures for the normal and Down's syndrome subjects were identical. The normal subjects were tested at Haskins Laboratories and the Down's syndrome subjects were tested at Celentano School over a four-day period. Each subject was scheduled for a 15-minute session each day. As a consequence of the difference in venue, Down's syndrome subjects' movements were analyzed via visicorder tracings (Honeywell, 1508C) while normal subjects' movements were fed directly into a PDP 11/45 for later computer analysis. For both types of recordings, a set number of degrees of movement corresponded to a calibrated voltage.

Each subject sat comfortably facing the apparatus with his right finger placed securely into the cylinder attached to the arm. Upon a signal from the experimenter, the subject moved to a specified steady-state joint angle (target) acting against a load supplied by the torque motor. The target angles for Experiment 1 were as follows: S1 = 155 deg, S2 = 165 deg, and S3 = 175 deg. When the subject's finger pointed straight ahead, the position was equal to 180 deg. The subject reached the target angle with a flexion movement from a starting position of 192 deg. The target angle was set by the subject by matching a line on the oscilloscope. When the angle was set, the subject closed his eyes and the oscilloscope was turned away. Within 0.5 to 2.0 sec after the target angle was achieved, a proportion of the load was released. The subject was instructed to maintain a steady resistance against the load--as indexed by the stable achievement of initial angle--and not to interfere voluntarily with the movement of the finger if the load was released.

For this study, partial unloading was achieved in the following manner. The control was set in servo-torque mode at the specified servo-position corresponding to the starting position as noted above. After the subject moved his finger to a specified joint angle, counteracting the resistance (100% torque) exactly, part of the load was released by switching from servo-torque control to torque control. This was achieved by the experimenter moving the manual switch on the control box. The amount of load released was regulated by setting the percentage of torque load. In order to obtain negative load release, the torque load was set at 5 and 10 percent but the direction of the resistance was reversed to act in the direction of flexion (i.e., in the direction that the finger was moving).

The measurements taken were changes in joint angle and changes in the amount of resistance acting on the finger. Measurements of joint angle change were taken by hand from visicorder tracings of finger displacement and measured to the nearest .5 cm, or obtained directly following A-to-D conversion (200 Hz) on the PDP 11/45. The criterion for determining the new steady-state joint angle was the point at which movement ceased (as shown by the movement tracings) for at least 500 msec (.5 cm on the tracing) following unloading.

A series of partial unloadings was conducted with each subject at each of three initial joint angles, S1, S2, and S3. For each series, seven separate unloadings were conducted (60%, 40%, 25%, 10%, 0%, -5%, and -10%), representing a percentage of maximum torque. The seven separate unloadings provided a sufficient amount of data to describe the torque by joint angle functions.

For each of the unloadings at least four trials were carried out. In all, more than 184 trials were conducted with each subject. The order in which the trials were given within each series, for the different unloadings, was randomly assigned. Series 1, 2, and 3 were presented to each subject in a predetermined balanced order.

Design

The data obtained from both Down's syndrome and normal subjects were graphed to determine if the torque by joint angle functions were parallel and non-intersecting as might be expected from the static equation, $F = -K\ (l - l_0)$. The algebraic error data were analyzed using a 2 (groups) x 3 (series) x 7 (unloadings) analysis of variance with repeated measures on the last two factors. A similar analysis was performed on variable error. In addition, tests of linearity (Wine, 1964) were also performed on the curves.

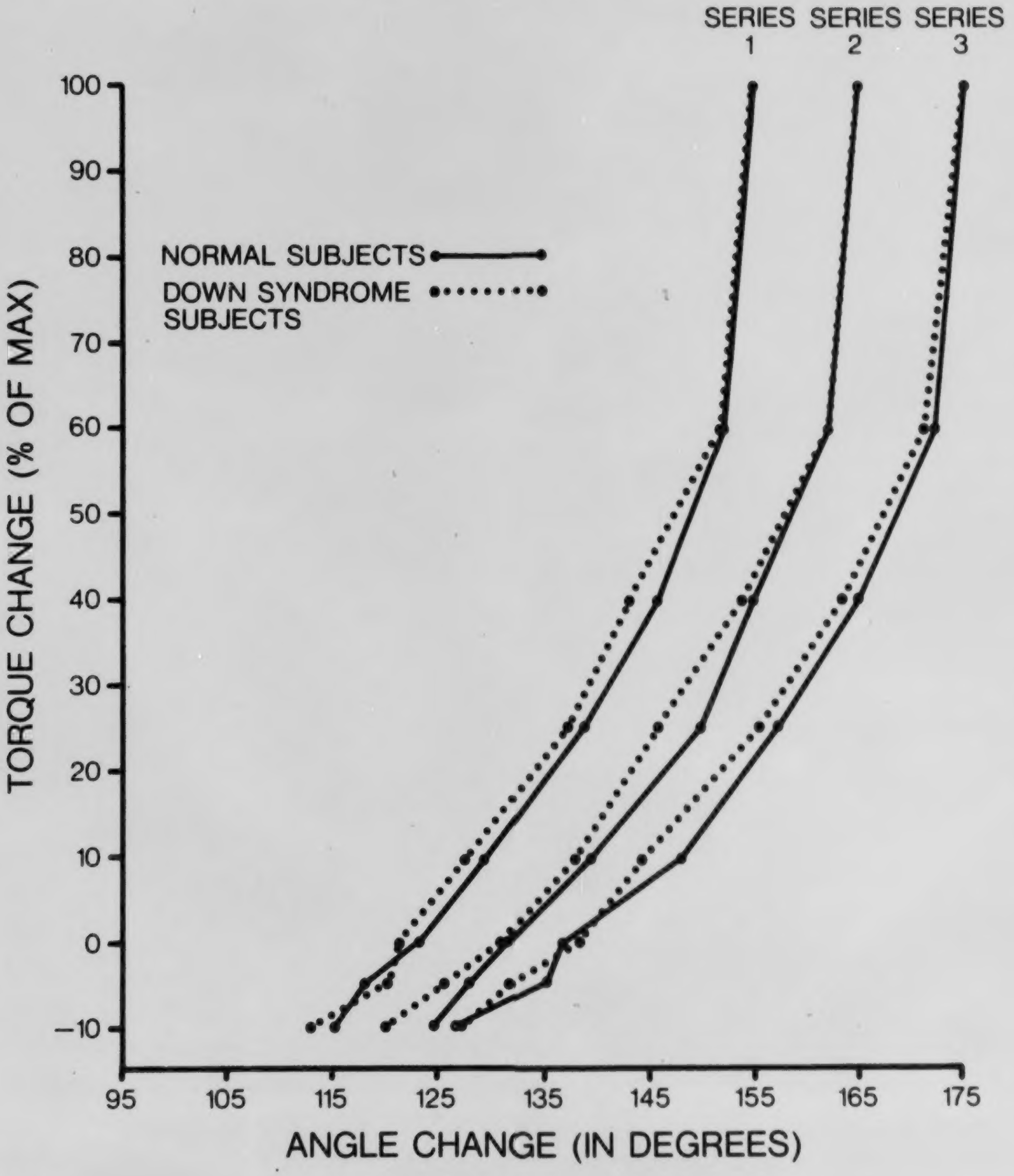
Results

Three sets of torque by joint angle functions were obtained as a result of partial unloading. The joint angle change in this case represents a deviation from the initial angle and conforms to standard measures of mean algebraic (constant) error. The results for the normal and Down's syndrome subjects are shown in Figure 1. The curves obtained for both normal and Down's syndrome subjects are indeed parallel and non-intersecting. It is clear from the figure that the two groups are nearly identical, and this was borne out by statistical analysis, $F\ (1,11) = .20$, $p > .05$. As expected, there was a significant torque change effect, $F\ (6,132) = 173.8$, $p < .001$, as well as a significant series effect, $F\ (2,132) = 6.23$, $p < .01$. There was also a significant series by torque interaction, $F\ (12,132) = 2.65$, $p < .01$. Analysis of this interaction revealed that for the three greatest unloadings (0%, -5%, and -10%), the joint angle change was larger between series 3 and 2 as well as series 3 and 1. Series 2 also revealed larger joint angle changes than series 1 at these unloading values.

Analysis of variable error, like constant error, reflected a high degree of similarity between Down's syndrome and normal groups. Only the torque effect was significant, $F\ (6,132) = 25.67$, $p < .001$. For the seven different unloadings (from 60% to -10%) variability in angle change appeared to increase systematically (1.0, 1.7, 2.4, 3.2, 3.8, 4.4, and 4.1 deg). No other main effect or interaction was significant for variable error.

A trend analysis of the F by l function obtained from each of the three series of unloadings for Down's syndrome and normal subjects was performed. The results revealed that the functions for both groups were essentially linear. For example, for the Down's syndrome group the proportion of variance accounted for by linearity in series 1, 2, and 3 was 94%, 93%, and 92%, respectively.

FIGURE 1. ANGLE CHANGE AS A FUNCTION OF TORQUE CHANGE
EXPERIMENT 1



Examination of the individual movement tracings, however, revealed some interesting differences. Representative movement tracings of individual subjects are presented in Figure 2A (normal subject SK) and 2B (Down's syndrome subject BV). The three movement tracings for each subject were recorded during unloadings of 60%, 10%, and -10% in Experiment 1. Qualitatively speaking, movements of normal subjects from the initial position to the target angle were, on the whole, more direct than the Down's syndrome subjects whose movements were more step-like or discontinuous (e.g., Brooks, 1974) and less stable. Although movement speed was not specified to the subjects beforehand, a rather interesting finding is that Down's syndrome subjects took significantly longer to reach the target angle. A random sample of 136 trials from Down's syndrome subjects was compared to the same number of trials for normal subjects. The mean movement time for the Down's syndrome group was 2.0 sec and the standard deviation was .4 sec compared to the mean and standard deviation of .8 sec and .1 sec for normal subjects. These movement times were significantly different from each other ($p < .001$).

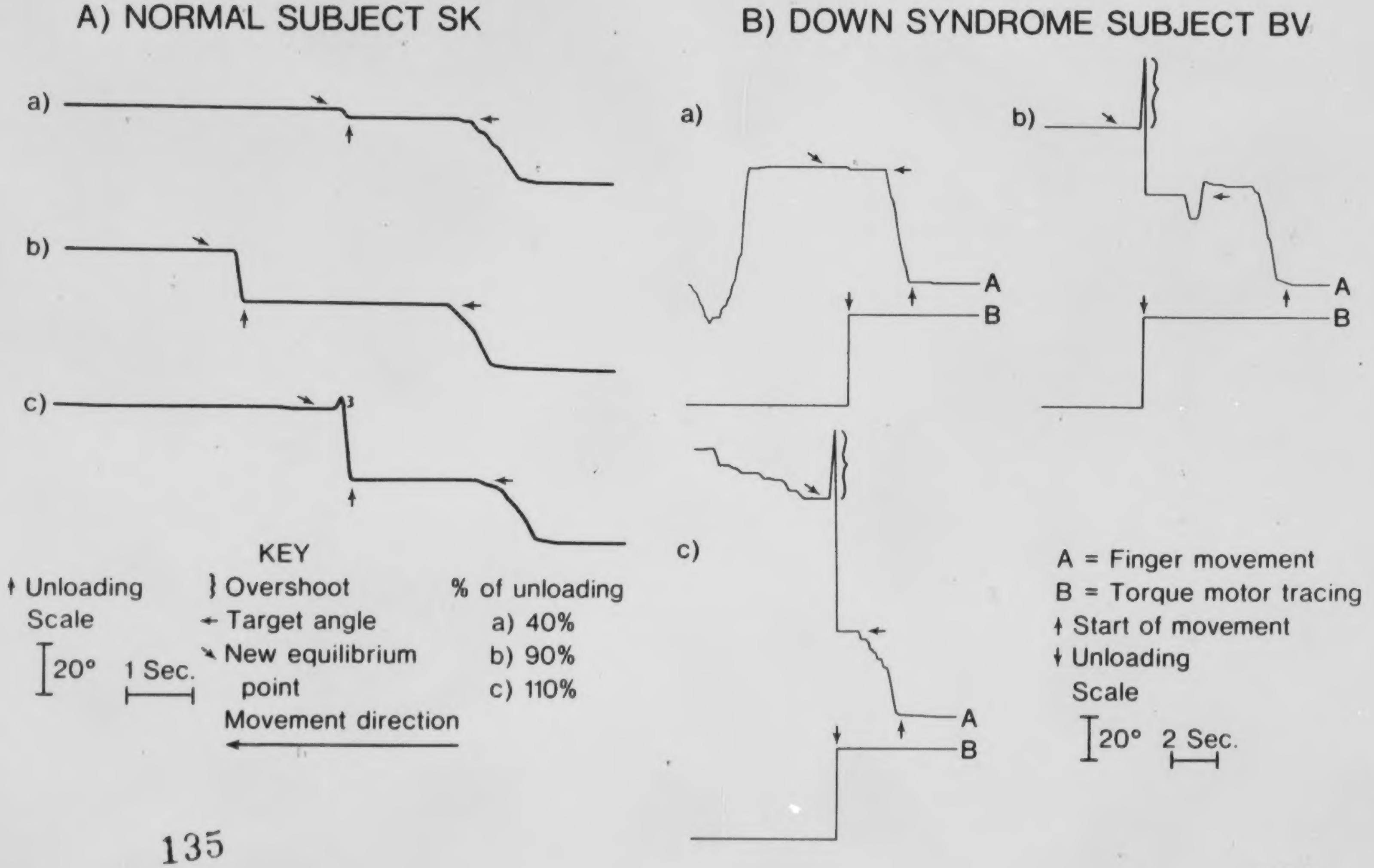
Down's syndrome subjects also differed from their normal peers in terms of oscillatory behavior at the newly established equilibrium position (i.e., following unloading, see Figure 2). Overshoot was measured to the nearest .5 degree for each of the trials and was analyzed using a group by series by torque analysis of variance with repeated measures on the last two factors. As shown in Table 1 a significantly greater amount of overshoot was found for Down's syndrome than for normal subjects, $\underline{F}\ (1,11) = 21.38$, $\underline{p} < .001$. There was also a significant series main effect, $\underline{F}\ (2,132) = 6.32$, $\underline{p} < .01$ and torque main effect, $\underline{F}\ (6,132) = 58.67$, $\underline{p} < .001$. The mean overshoot for each series was 4.87, 5.48, and 6.91 deg, respectively. As the amount of unloading increased (and thus angle change increased), the amount of overshoot increased. This finding holds for both groups but is magnified in the Down's syndrome group as evident in a group by torque interaction, $\underline{F}\ (1,132) = 17.27$, $\underline{p} < .001$ (see Table 1).

Discussion

The results of Experiment 1 support the notion that when muscles are constrained to act as a unit in controlling movement about a joint, that unit behaves qualitatively like a mass-spring system. The three sets of torque by joint angle functions obtained for both normal and Down's syndrome subjects were parallel and non-intersecting, and concur with the findings of Asatryan and Fel'dman (1965). It may be reasoned that, for the subjects in this study, the $-K$ and l_0 parameters were established in counterbalancing the external force to maintain the specified joint angle. A systematic angle change accompanied the systematic torque change. Thus, it appears that the parameters of $-K$ and l_0 remained relatively constant during unloading. Furthermore, when the subjects were asked, on different occasions, to reach new joint angles, new zero angles were established. The change in zero length resulted in parallel and non-intersecting functions.

Of course the significant result of the present experiment is that the torque by joint angle functions for Down's syndrome subjects and a normal group were practically identical (see Figure 1). It appears therefore that the underlying movement organization, at least under static conditions, is

FIGURE 2. TRACINGS OF FINGER MOVEMENT
EXPERIMENT 1



basically similar in both populations. From a mass-spring perspective, Down's syndrome subjects, like normals, appear capable of specifying system parameters--stiffness and equilibrium length--that may determine movement at the joint.

Although the curves appear graphically non-linear, the torque by joint angle functions were characterized by a statistically linear trend. Thus, the non-linear component in this study is somewhat less than might be expected based on the findings and qualitative analysis of Asatryan and Fel'dman (1965). However, these authors did not subject their obtained functions to any statistical analysis as we have done here.

One way of interpreting the present findings is that the nervous system (in both Down's syndrome and normal populations) is able to "tune" the muscle-joint system by adjusting the length-tension relationships of the muscles involved. In the simple case, agonist-antagonist pairs can be represented by parallel length-tension curves whose slopes correspond to muscle stiffness. A change in innervation rate to one muscle or the other will shift the equilibrium point of the system (cf. Bizzi, 1980; Cooke, 1980; Fel'dman, 1966b, 1980a, 1980b; Kelso & Holt, 1980). In this view, supported by the ability of deafferented monkeys (Bizzi et al., 1978; Bizzi et al., 1976) and humans without intact joint and cutaneous reception (Kelso, 1977; Kelso & Holt, 1980) to localize limbs accurately, stiffness is set prior to movement and is a control parameter. On the other hand, Houk (1978) has presented evidence in favor of a view in which a combination of muscle spindle and tendon proprioceptors provides feedback about muscle stiffness. In this view, stiffness is a regulated variable of the system. Regardless of which view one adopts, both are consonant with the perspective offered here (but see the General Discussion for possible qualifications on this view, and also Cooke, 1980, for a model of how mechanical and reflex variables may interact). It is the specification of dynamic variables (e.g., stiffness, damping) rather than kinematic variables (e.g., displacement, velocity) that appropriately characterize the neuromuscular organization of the muscle-joint system.

There were, however, qualitative differences among the movement patterns of Down's syndrome and normal subjects. There were clear differences between the graphs in the trajectories toward the target angle (see Figure 2). Moreover, Down's syndrome subjects were less able, after reaching the target, to maintain a steady position. One possible explanation for the latter finding is that once the target was reached, visual guidance was removed. In a previous study (Davis & Kelso, Note 1), it was found that Down's syndrome subjects were less able than normal subjects to reproduce movements accurately without visual guidance. However, the movements of Down's syndrome subjects in the present study were also less smooth and accurate when visual guidance was available. In the present study, movements were made by matching a cursor to a fixed line target on an oscilloscope screen. Visual guidance from an oscilloscope may not be the same as direct visual guidance of the finger. Nevertheless, under the present conditions of visual guidance to a target as well as maintaining a set joint angle with visual information absent, Down's syndrome subjects were not as accurate as their normal peers.

Second, significantly greater overshoot or fluctuation about the equilibrium point was found in Down's syndrome subjects following unloading.

Table 1

Mean Overshoot (in degrees) for Each Group as a Function of Torque Loading

	% of Torque Unloading													
	60%		40%		25%		10%		0%		-5%		-10%	
	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
Experiment 1														
Down Syndrome	0.2	0.3	1.6	1.6	4.2	3.0	9.9	4.3	14.3	5.4	17.1	6.2	18.0	5.2
Normal	0.1	0.1	0.1	0.1	0.7	0.5	1.6	1.1	3.3	1.6	4.2	2.1	5.4	2.2
Experiment 2														
Down Syndrome	0.1	0.1	--		3.5	2.3	--		9.4	4.9	--		12.5	5.5
Normal	0.2	0.3	--		0.9	0.5	--		1.3	1.1	--		1.1	1.9

Overshoot may be taken as an index of the damping parameter. For example, an underdamped system will fluctuate about the equilibrium position and, for most purposes, is unstable. On the other hand, an overdamped system exhibits slowed movement speed and no oscillation about the equilibrium point. A critically damped system is one in which the movement will reach the equilibrium position in the fastest possible time. Most human muscle joint systems appear to be damped just under critical (cf. Kelso & Holt, 1980; Neilson & Neilson, 1978). Neilson and Neilson (1978), for example, found their normal subjects did not exceed five percent of the movement arc in overshoot during a rapid voluntary movement.

The results of our Experiment 1 conform to the above figure; overshoot was found to be 8.4% of the total movement arc for all normal subjects' movements. Different amounts of overshoot may be due to the differences in the movement conditions between the present study and that of Neilson and Neilson (1978). Overshoot was measured in the Neilson and Neilson study following a rapid voluntary movement. In our study, however, overshoot was measured following unloading during which an active halting of the limb was assumed not to have occurred. It seems reasonable, therefore, to expect some difference in overshoot between the normal subjects in our study and those in the Neilson and Neilson (1978) study. Perhaps the more important finding is that Down's syndrome subjects, in sharp contrast to normal subjects, appear to behave in an underdamped manner, as suggested by the 27.4% of overshoot found for this group.

Although the results of Experiment 1 revealed that the torque by joint angle functions were near-identical for Down's syndrome and normal subjects, a question remains as to whether both populations can alter the slope of the functions (specify stiffness) to a similar degree. Following Asatryan and Fel'dman (1965), one way to examine this question is to require subjects to tense the muscles voluntarily prior to unloading. "Stiffening" the muscle-joint system in this manner should reduce the amount of absolute joint-angle change, thus increasing the slope of the torque-joint angle functions. In Experiment 2, the foregoing strategy was employed to determine whether Down's syndrome individuals could control stiffness as effectively as normal subjects.

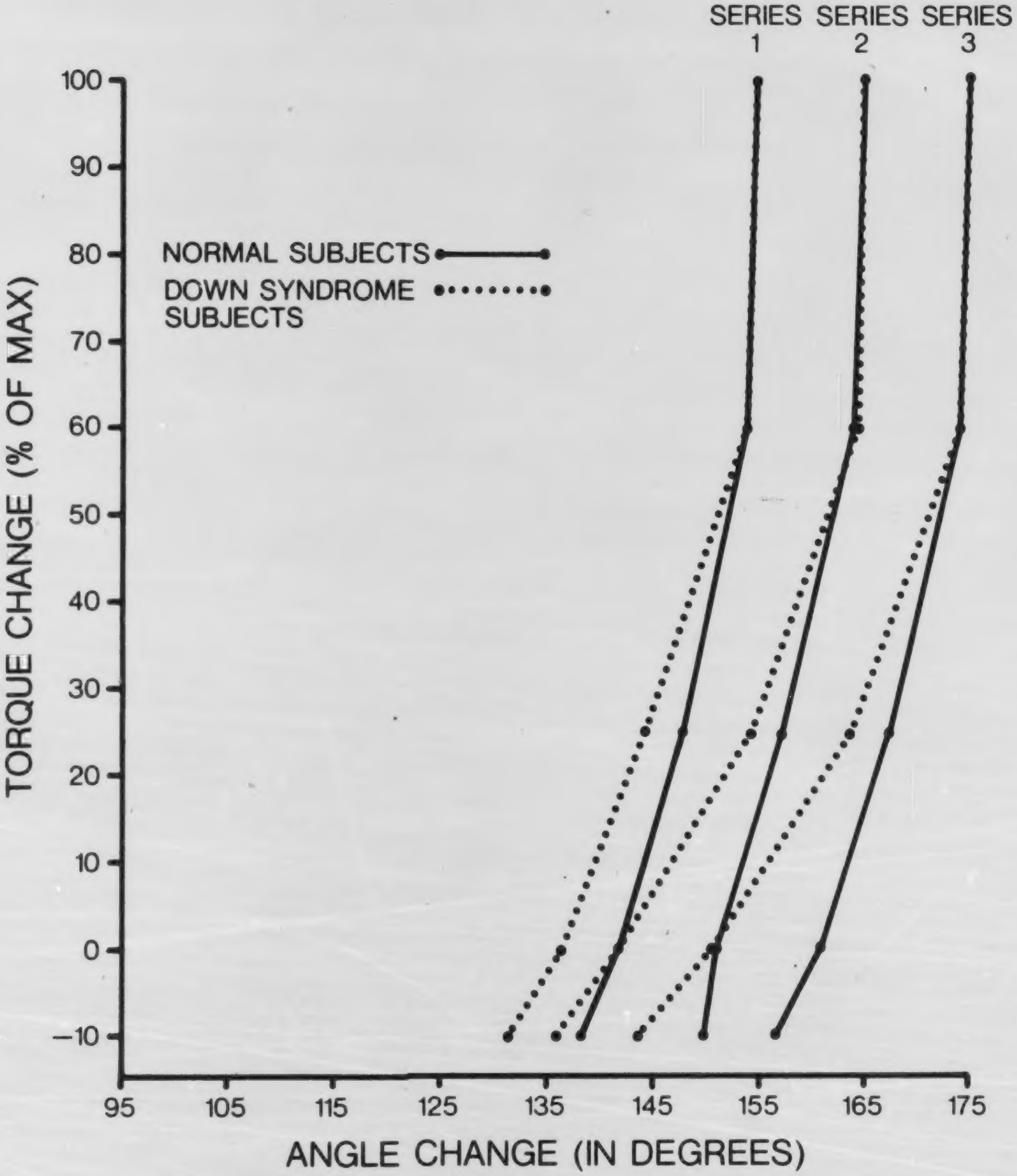
EXPERIMENT 2

Method

The subjects, methods, and procedures in Experiment 2 were the same as in Experiment 1 except that subjects were instructed to tense (cocontract) their muscles in an effort to maintain the joint angle against the perturbation. Each subject was given some practice prior to the experiment proper to ensure that the instructions were understood. In the experimental trials, on reaching the target angle, subjects were asked to "stiffen their muscles" in order to maintain the joint angle.

Only four separate unloadings (60%, 25%, 0%, and -10%) were used for Experiment 2. These four unloadings provided a sufficient description of the

FIGURE 3. ANGLE CHANGE AS A FUNCTION OF TORQUE CHANGE
EXPERIMENT 2



torque by joint angle function without inducing undue fatigue. For each of the unloadings, at least four trials were carried out and more than 48 trials were conducted for each subject.

Data were graphed as in Experiment 1 and analyzed as before using a 2 (group) x 3 (series) x 4 (unloadings) analysis of variance with repeated measures on the last two factors.

Results

Changes in joint angle under "muscle stiffening" conditions are shown in Figure 3 for normal and Down's syndrome subjects. Visual inspection of the figure suggests some potential difference in the torque-joint angle functions of normal and Down's syndrome subjects. Although overall differences between the groups just failed to reach significance, F (1,11) = 4.35, $p > .05$, there was a significant torque effect, F (3,66) = 69.86, $p < .001$, and group by torque interaction, F (3,66) = 4.61, $p < .01$. Inspection of the means shows that joint angle increases as a function of the magnitude of unloading, and that this effect is greater for Down's syndrome subjects (see Table 1). The mean angle changes for each unloading for Down's syndrome subjects were 1.42, 12.40, 23.49, and 29.52 deg and 1.29, 8.08, 14.66, and 17.84 deg for normal subjects.

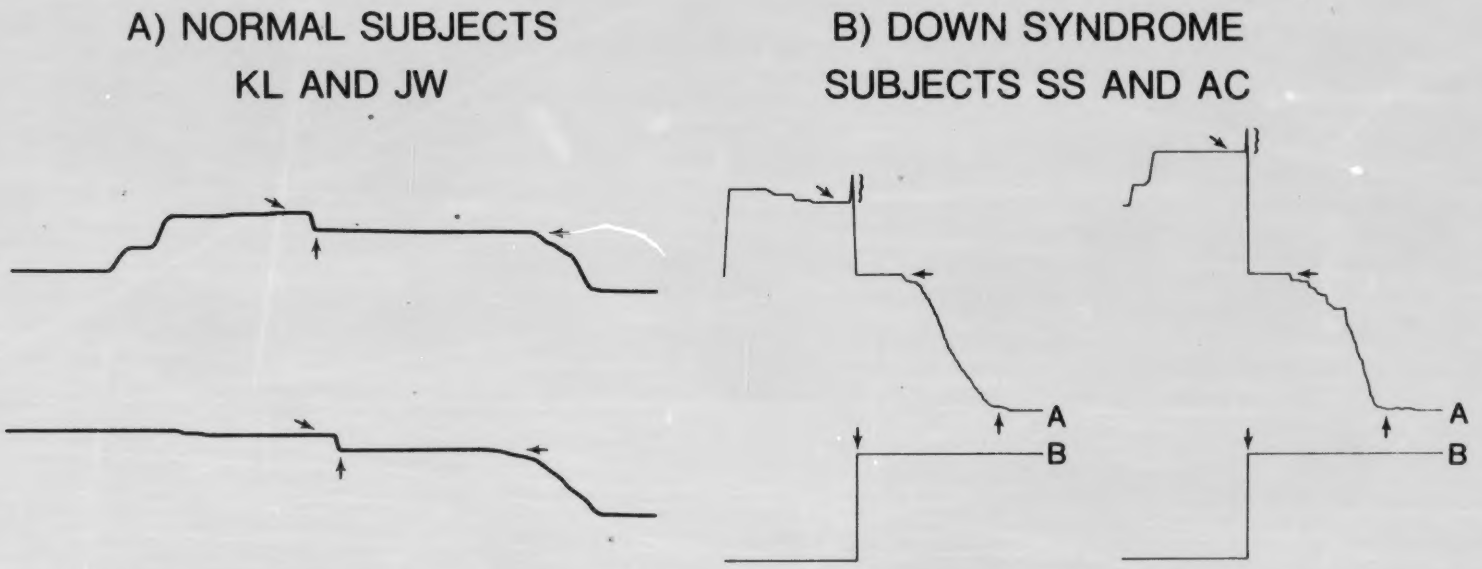
A series by torque interaction was also found, F (3,66) = 3.13, $p < .01$. As in Experiment 1, at the higher levels of unloading (the last two in Experiment 2) the magnitude of angle change increased between series 1 and 2 and between 2 and 3.

There were no significant differences between the groups in the standard deviation of mean angle change, F (1,11) = 3.18, $p > .05$. However, there was a significant torque effect, F (3,66) = 20.62, $p < .001$ and a significant torque by series interaction, F (6,66) = 4.76, $p < .001$. Variability increased as the magnitude of the unloading increased but was significant only in series 1 and 3.

Representative movement tracings from Down's syndrome and normal subjects in Experiment 2 are shown in Figure 4. Qualitatively, the movements of the subjects are similar to those in Experiment 1. Down's syndrome subjects appeared less capable of reaching and maintaining the target angle (compare Figures 2 and 4).

Likewise, overshoot or oscillation about the equilibrium point following unloading was amplified in Down's syndrome subjects (see Table 1) and was revealed in a group main effect, F (1,11) = 8.88, $p < .05$. The main effects of series, F (2,66) = 5.96, $p < .01$ and torque, F (3,66) = 1.82, $p < .001$, were also significant. Increases in overshoot occurred from series 1 to 2 and from 2 to 3, and were also evident as the magnitude of unloading increased. Increases in overshoot held for each series in Down's syndrome subjects but not normals, and resulted in a significant group x series interaction, F (2,66) = 4.39, $p < .05$. Significant group x torque, F (3,66) = 8.3, $p < .001$, and series by torque, F (6,66) = 3.95, $p < .01$, interactions were also found. As the unloading increased, the degree of overshoot increased; this effect was

FIGURE 4. TRACINGS OF THE FINGER MOVEMENT EXPERIMENT 2


A) NORMAL SUBJECTS KL AND JW

B) DOWN SYNDROME SUBJECTS SS AND AC

KEY

↑ Unloading
Scale
20° 1 Sec.

} Overshoot
← Target angle
↘ New equilibrium point
Movement direction ←
% of unloading 110%

A = Finger movement
B = Torque motor tracing
↑ Start of movement
↓ Unloading
Scale
20° 2 Sec.

magnified for Down's syndrome subjects and for series 2 and 3 (see Table 1, Experiment 2).

Discussion

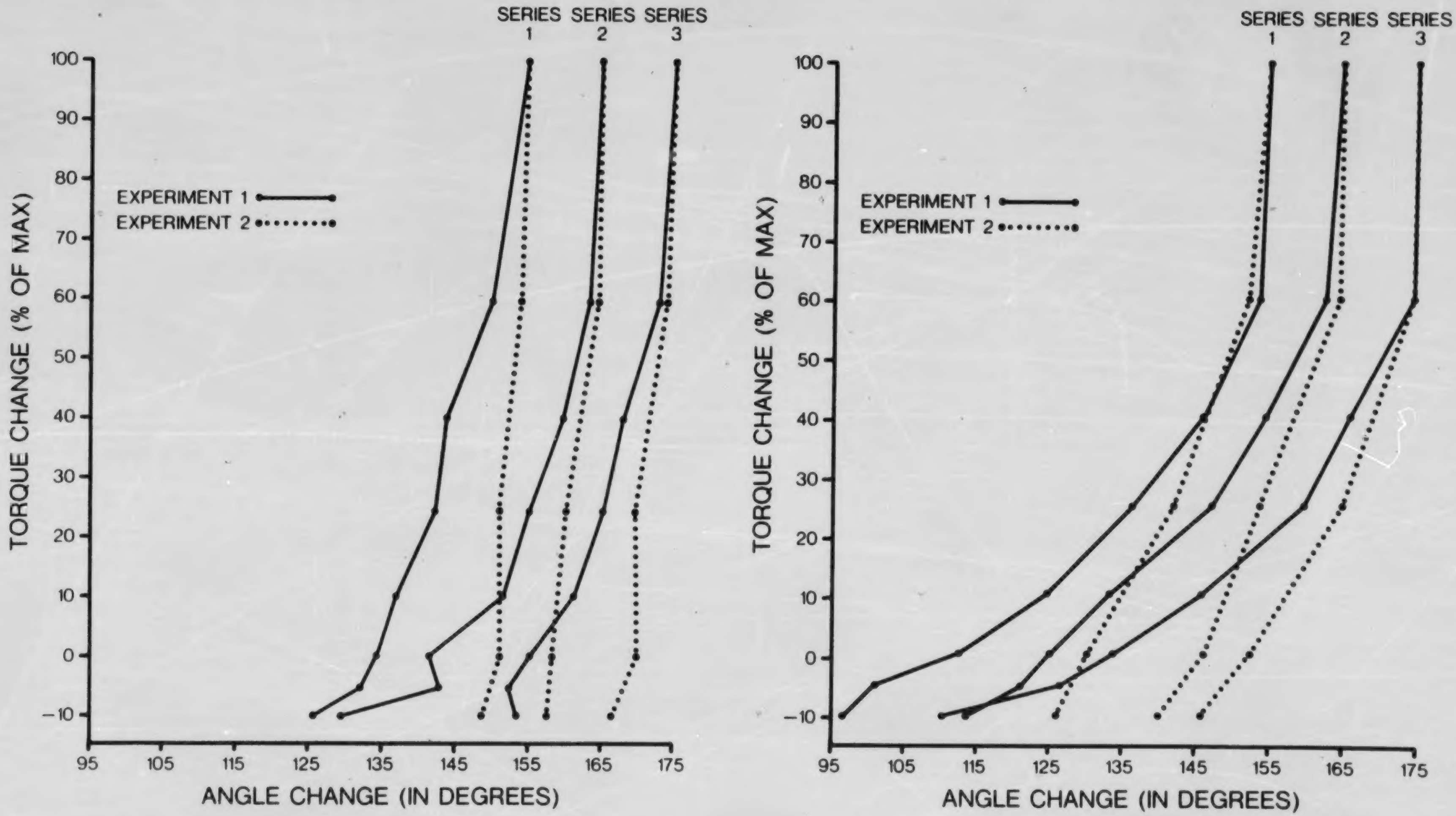
In the second experiment it was demonstrated that Down's syndrome subjects were able to increase the stiffness parameter voluntarily when asked to tense their muscles against the load change. This capability is shown by the increase in slope of the torque by joint angle functions from Experiment 1 to Experiment 2 (compare Figures 1 and 3). It is well known that an increase in activation of motoneurons increases the stiffness property of the muscle (Agarwal & Gottlieb, 1977; Andreeva & Shafranova, 1975; Barmack, 1976; Houk, Singer, & Goldman, 1970; Rack, 1969; Safronov, 1970). An increase in stiffness can occur without an increase in force or without a change in joint angle.

Normal subjects were also capable of increasing stiffness and apparently to a somewhat greater extent than Down's syndrome subjects. The mean slopes of the functions generated in both experiments were 1.77 and 3.56, respectively, for Down's syndrome subjects and 1.83 and 4.67 for normal subjects. There was no statistically significant difference between groups on overall angle change in Experiment 2; however, at higher levels of torque change, significant differences between groups did exist as revealed in the significant group by torque interaction.

There were also some noteworthy differences among individual subjects that warrant some discussion. For example, Down's syndrome subject LM's stiffness characteristics were closer to that found for normal subjects (see Figure 5) and, in fact, exceeded some of the normal subjects. Likewise, one normal subject, GA, displayed muscle stiffness more aligned to the Down's syndrome group than to the normal group (see Figure 5). If these extremes are excluded from the groups, the difference between Down's syndrome and normal subjects is magnified. That wide individual differences among subjects exist within both groups is expected. It is known that stiffness varies between and within normal individuals (Safronov, 1970). Also, extreme individual differences among Down's syndrome subjects on a number of psychological and physical variables have been found (James, 1974; LaVeck & Brehm, 1978). It can be reasoned that stiffness, as with many other variables, operates on a continuum rather than strictly dichotomizing the populations observed.

Finally, there is some indirect evidence that suggests Down's syndrome subjects have reduced capacity of muscle activation that might be associated with the ability to regulate muscle stiffness. Most Down's syndrome infants are deficient in the amino acid 5-hydroxytryptophan (Coleman, 1973; Koch & de la Cruz, 1975), which is thought to play an important role in neural transmission (McCoy, Segal, & Strynadka, 1975) and muscle contraction (Ahlman, Grillner, & Udo, 1971). The finding that Down's syndrome subjects have slower movement response times (Lange, 1970) may support the idea that Down's syndrome subjects are less able to activate their muscles. If, as some have claimed (cf. Lestienne, Polit, & Bizzi, 1980), integrated electromyography reflects active stiffness in muscles, then it seems worthwhile to subject the foregoing speculations regarding stiffness to further experimental test.

FIGURE 5. ANGLE CHANGE AS A FUNCTION OF TORQUE CHANGE



A) DOWN SYNDROME SUBJECT LM

B) NORMAL SUBJECT GA

GENERAL DISCUSSION

A major finding in this study is that the gross underlying organization of motor control in Down's syndrome subjects is qualitatively similar to normal subjects and can be described, to a first approximation, in terms of a vibratory system. We have found that the torque by joint angle curves, described as invariant characteristics of the system by Asatryan and Fel'dman (1965) (see also Fel'dman, 1980a, 1980b), are obtainable from Down's syndrome subjects under two conditions of maintaining a steady joint angle against a load. In one condition the subjects did not voluntarily intervene during partial unloading. In the other condition the subjects voluntarily tensed their muscles, which resulted in an increase in the active stiffness of the system. The importance of obtaining the invariant characteristics (static muscle torque versus angle) in this special Down's syndrome population is magnified by the recent findings of Fel'dman (1980a, 1980b). With normal subjects, Fel'dman has shown that the Invariant Characteristic (IC) may characterize the behavior of muscle-joint systems not only during the maintenance of a steady posture (Asatryan & Fel'dman, 1965) but also during rhythmic (Fel'dman, 1980a) and discrete movements (Fel'dman, 1980b). For example, in order to maintain a steady angle against changing loads, the system need only shift from one IC to another. Referring to Figure 1, it can be seen that in order to maintain a 150 deg joint angle against a 60% load, the IC of series 1 is used. In order to maintain the 150 deg angle when the load changes to 25%, a shift from the IC of series 1 to the IC of series 2 is required. This transition from one IC to another appears to be effected through a change of the threshold angle at which motor units are recruited (Fel'dman, 1966a, 1980b; see also Crago, Houk, & Hasan, 1976).

By the same token, movements may be accomplished by shifts along the form of the invariant characteristics, that is, by shifts of the equilibrium point of the muscle load system and by changes in the form of the IC (Fel'dman, 1974a, 1974b). The latter is shown by the set of ICs obtained in Experiment 1 (Figure 1) and the set obtained in Experiment 2 (Figure 3). Through cocontraction of the antagonist muscles, the stiffness of the muscle load system may be increased and this is associated with increases in slope of the IC. It may be assumed that during movement, transformation from one set of curves to another is possible (Fel'dman, 1980a, 1980b). Thus, movements may be achieved through simple changes in the parameters of the muscle-load system. According to this view, amplitude of movement (position) may be regulated through changes in zero length, and velocity and acceleration through changes in stiffness (cf. Kelso & Holt, 1980).

A cautionary point worth emphasizing here is that the majority of experiments (including ours), and their consequent interpretation, deal with movements in a very restrained environment (e.g., sitting down with shoulder or wrist position fixed). New data reveal that the pattern of stiffness changes at a joint (say, the arm) is mutable, depending in a significant way on the postural status of the subject (Nashner, Note 2). Some modification (or even rejection) of the type of model proposed here for rather fixed actions at a joint may well be in order when more real-life situations are examined (e.g., a standing subject supported to varying degrees).

Although there appear to be overall similarities in the gross organization of motor control of Down's syndrome and normal subjects, there are notable differences in the precision with which Down's syndrome subjects attain target positions (see Figures 2 and 4). As shown by individual movement tracings and subsequent analysis, the movement patterns were qualitatively different between Down's syndrome and normal subjects in both Experiments 1 and 2. The findings here are consistent with other studies that show Down's syndrome subjects to be less accurate in controlling movements than their normal peers (e.g., Davis & Kelso, Note 1).

Relatedly, and perhaps most important, Down's syndrome subjects differed from their normal peers in oscillatory behavior about the newly established equilibrium position (i.e., following unloading, see Figures 2 and 4). Oscillatory behavior is taken as an index of the damping parameter. As previously noted, the finding that underdamping characterizes the muscle-joint system of Down's syndrome subjects is consistent with the finding that these subjects are less accurate in movement than normal subjects. It is the underdamping characteristic, in addition to stiffness, that may distinguish Down's syndrome from normal subjects. In this regard, it is important to note that although individual subjects GA (normal) and LM (Down's syndrome) were unlike their respective groups with respect to stiffness, they did not deviate from group performance on measures of damping. Mean overshoot for GA was 2.56 deg compared to the overall mean of 2.2 for the normal group. Likewise the means for LM and the Down's syndrome group were 9.35 and 9.31, respectively. It is not known whether Down's syndrome subjects are able to modify the damping parameter through training. Other studies have suggested that normal subjects can be trained to regulate damping during voluntary movement (e.g., Neilson & Lance, 1978). But further investigations are needed to determine to what extent Down's syndrome subjects have this capacity.

Overall, a number of findings concerning the motor control of Down's syndrome subjects are provided in this study and several avenues of research are suggested. We found that the Down's syndrome subjects in this study are not readily distinguishable from normal subjects in terms of gross movement organization; systematic torque by joint angle functions were obtained for both groups. Apparently, muscles are constrained to act as a unit in both normal and Down's syndrome subjects and this unit exhibits behavior--to a first approximation--qualitatively similar to a mass-spring system (cf. Kelso, Holt, Kugler, & Turvey, 1980, for review). Where the groups differ, however, is in the specification of stiffness, especially at high values of torque unloading and in the damping characteristic. It is interesting in this regard that a recent analog model--similar to the one under consideration here--characterizes hypotonia in terms of decreases in resting stiffness (Cooke, 1980). If our interpretation is reasonable, we may advance the hypothesis that it is a deficiency in setting damping and stiffness parameters that best characterizes the motor behavior of people with Down's syndrome--at least in simple, discrete movements. This view promotes a trend away from more descriptive terms, like hypotonia, that have been used up to now.

REFERENCE NOTES

1. Davis, W. E. & Kelso, J. A. S. *Limb localization in Down's syndrome subjects*. Unpublished study, Haskins Laboratories, 1979.
2. Nashner, L. M. *Central and peripheral components of a model for the coordination of balance*. Paper presented at the Sloane Foundation Conference on Coordination of Action, University of California, San Diego, February, 1981.

REFERENCES

Agarwal, G. C., & Gottlieb, G. L. Oscillation of the human ankle joint in response to applied sinusoidal torque on the foot. *Journal of Physiology*, 1977, *268*, 151-176.

Ahlman, H., Grillner, S., & Udo, M. The effect of 5-HTP on the static fusimotor activity and the tonic stretch reflex of an extensor muscle. *Brain Research*, 1971, *27*, 393-396.

Andreeva, E. A., & Shafranova, E. I. Muscular activity control systems for Parkinson's disease. *Automation and Remote Control*, 1975, *36*, 596-605.

Asatryan, D. G., & Fel'dman, A. G. Functional tuning of the nervous system with control of movement or maintenance of a steady posture - I. Mechanographic analysis on the work of the joint on execution of a postural task. *Biophysics*, 1965, *10*, 925-935.

Barmack, N. H. Measurements of stiffness of extraocular muscle of the rabbit. *Journal of Neurophysiology*, 1976, *39*, 1009-1019.

Berkson, G. An analysis of reaction time in normal and mentally deficient young men, III. Variation of stimulus and of response complexity. *Journal of Mental Deficiency Research*, 1960, *4*, 69-77.

Bizzi, E. Central and peripheral mechanisms in motor control. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior*. Amsterdam: North-Holland Publishing Company, 1980.

Bizzi, E., Dev, P., Morasso, P., & Polit, A. Effects of load disturbance during centrally initiated movements. *Journal of Neurophysiology*, 1978, *41*, 542-556.

Bizzi, E., Polit, A., & Morasso, P. Mechanisms underlying achievement of final head position. *Journal of Neurophysiology*, 1976, *39*, 435-444.

Brooks, V. B. Some examples of programmed limb movements. *Brain Research*, 1974, *71*, 299-308.

Coleman, M. *Serotonin in Down syndrome*. New York: American Elsevier Publishing Company, 1973.

Cooke, J. D. The organization of simple, skilled movements. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior*. Amsterdam: North-Holland Publishing Company, 1980.

Cowie, V. A. *A study of the early development of Mongols*. Oxford: Pergamon Press, 1970.

Crago, P. E., Houk, J. C., & Hasan, Z. Regulatory action of the human stretch reflex. *Journal of Neurophysiology*, 1976, *39*, 925-935.

Fel'dman, A. G. Functional tuning of the nervous system with control of movement or maintenance of a steady posture-II. Controllable parameters of the muscle. *Biophysics*, 1966, *11*, 766-775. (a)

Fel'dman, A. G. Functional tuning of the nervous system with control of movement or maintenance of a steady posture-III. Mechanographic analysis of execution by man of the simplest motor tasks. Biophysics, 1966, 11, 766-775. (b)

Fel'dman, A. G. Changes in the length of the muscle as a consequence of a shift in equilibrium in the muscle-load system. Biophysics, 1974, 19, 534-538. (a)

Fel'dman, A. G. Control in the length of a muscle. Biophysics, 1974, 19, 766-711. (b)

Fel'dman, A. G. Superposition of motor programs. I. Rhythmic forearm movements in man. Neuroscience, 1980, 5, 81-90. (a)

Fel'dman, A. G. Superposition of motor programs--II. Rapid forearm flexion in man. Neuroscience, 1980, 5, 91-95. (b)

Frith, U., & Frith, C. D. Specific motor disabilities in Down syndrome. Journal of Child Psychology and Psychiatry, 1974, 15, 293-301.

Houk, J. D. Participation of reflex mechanisms and reaction time processes in compensatory adjustments to mechanical disturbances. In J. Desmedt (Ed.), Cerebral motor control in man: Long loop mechanisms. Basel: Karger, 1978.

Houk, J. C., Singer, J. J., & Goldman, M. R. An evaluation of length and force feedback to soleus muscle of decerebrate cats. Journal of Neurophysiology, 1970, 33, 784-811.

James, R. J. Multivariate analysis of the walking behavior in institutionalized Down's syndrome males. Unpublished doctoral dissertation, University of Wisconsin, 1974.

Kelso, J. A. S. Motor control mechanisms underlying human movement reproduction. Journal of Experimental Psychology, 1977, 8, 429-443.

Kelso, J. A. S., & Holt, K. G. Exploring a vibratory systems analysis of human movement production. Journal of Neurophysiology, 1980, 43, 1183-1196.

Kelso, J. A. S., Holt, K. G., & Flatt, A. E. The role of proprioception in the perception and control of human movement: Toward a theoretical reassessment. Perception & Psychophysics, 1980, 28, 45-52.

Kelso, J. A. S., Holt, K. G., Kugler, P. N., & Turvey, M. T. On the concept of coordinative structure as dissipative structure--II. Empirical lines of convergence. In G. E. Stelmach & J. Requin (Eds.), Tutorials in motor behavior. Amsterdam: North Holland, 1980.

Knight, R. M., Atkinson, B. R., & Hyman, J. A. Tactual discrimination and motor skills in mongoloid and non-mongoloid retarded and normal children. American Journal of Mental Deficiency, 1967, 71, 894-900.

Koch, R., & de la Cruz, F. F. Down's syndrome (Mongolism): Research, prevention and management. New York: Brunner/Mazel, 1975.

Lange, E. K. A comparison of the response times of mongoloid children and normal children. Unpublished doctoral dissertation, University of New Mexico, 1970.

LaVeck, B., & Brehm, S. S. Individual variation among children with Down syndrome. Mental Retardation, 1978, 16, 135-137.

Lestienne, F., Polit, A., & Bizzi, E. From movement to posture. In C. H. Nadeau, K. M. Newell, G. C. Roberts, & H. H. Halliwell (Eds.), The skillfulness in movement: Theory and application. Champaign, Ill.: Human Kinetics Press, 1980.

McCoy, E. E., Segal, D. J., & Strynadka, I. Decreased ATPase, increased NA+ and decreased K+ in Down's syndrome platelets. In R. Koch & F. F. de la Cruz (Eds.), Down's syndrome (Mongolism): Research, prevention, and management. New York: Brunner/Mazel, 1975.

Neilson, P. D., & Lance, J. W. Reflex transmission characteristics during voluntary activity in normal man and patients with movement disorders. In J. E. Desmedt (Ed.), *Cerebral motor control in man: Long loop mechanisms: Progress in clinical neurophysiology* (Vol. 4). Karger: Basel, 1978.

Neilson, P. D., & Neilson, M. D. The role of action reflexes in the damping of mechanical oscillations. *Brain Research*, 1978, *142*, 439-453.

Owens, D., Dawson, J. C., & Losin, S. Alzheimer's disease in Down's syndrome. *American Journal of Mental Deficiency*, 1971, *75*, 606-612.

Polit, A., & Bizzi, E. Processes controlling arm movements in monkeys. *Science*, 1978, *201*, 1235-1237.

Rack, P. M. H. The significance of mechanical properties of muscle in the reflex control of posture. In P. Anderson & A. K. S. Jansen (Eds.), *Excitatory synaptic mechanisms*. Bergen, Tromio: Universitetsforlaget, Oslo, 1969.

Safronov, V. A. Problem of the regulation of muscle tone. *Biophysics*, 1970, *15*, 1145-1154.

Schmidt, R. A., & McGown, C. Terminal accuracy of unexpectedly loaded rapid movements: Evidence for a mass-spring mechanism in programming. *Journal of Motor Behavior*, 1980, *12*, 149-161.

Turvey, M. T. Preliminaries to a theory of action with reference to vision. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting and knowing: Toward an ecological psychology*. Hillsdale, N.J.: Erlbaum, 1977.

von Bertalanffy, L. *General system theory*. London: Penguin University Books, 1973.

Wine, R. L. *Statistics for scientists and engineers*. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1964.

FOOTNOTES

[1] Note that equifinality can be disrupted in deafferented but not normal animals when the postural relations between animal and apparatus are changed (Polit & Bizzi, 1978). When the latter are fixed, however, deafferented monkeys and functionally deafferented humans (Kelso & Holt, 1980; Kelso, Holt, & Flatt, 1980) exhibit equifinality when the initial position of the limb is altered unexpectedly.

PERIODICITY AND AUDITORY MEMORY: A PILOT STUDY

Janet May and Bruno H. Repp

Abstract. Band-limited periodic and aperiodic nonspeech stimuli varying in center frequency but having similar spectral envelopes were presented in same-different tasks at two interstimulus intervals (.5 and 2 sec). Discriminability decreased as a function of interval, but there was no difference between the types of stimuli, suggesting that periodic and aperiodic stimuli are equally well retained in auditory memory.

Studies of categorical perception have sometimes found unexplained differences between stimulus sets. For example, May (1979) tested Egyptian listeners' discrimination of vowel-fricative-vowel syllables drawn from continua spanning two Arabic fricative contrasts, /ɣ-ʕ/ and /x-ħ/, that differ from each other only in voicing, having equivalent places of articulation. The subjects in that study found the stimuli from the voiced fricative continuum easier to discriminate than those from the voiceless continuum. Although the two continua differed along several acoustic dimensions, the most conspicuous of these was the periodicity present during the frication portions of the voiced continuum and absent on the voiceless continuum. The stimuli of another recent study of categorical perception (Healy & Repp, 1982) included vowels from an /i-I/ continuum and isolated fricative noises from a /s-ʃ/ continuum. When listeners were asked to categorize stimuli from either continuum presented in pairs, the contrastive effect exerted by one member of the pair on the labeling of the other member was much larger for vowels than for fricative noises, even when differences in discriminability were taken into account. Once again, presence versus absence of periodicity seemed the most salient feature distinguishing the two types of stimuli, although there were several other differences as well.

At least two reasons could be proposed why periodicity might make a difference in auditory memory. On one hand, it seems possible that the recurrence of signal portions with similar structure--i.e., of the successive pitch periods--strengthens the auditory trace, in the way repetitions of a stimulus generally do. According to this hypothesis, the auditory traces of aperiodic stimuli would not have the benefit of such reinforcement-by-repetition and, therefore, would be weaker than those of periodic stimuli. On the other hand, the repetitive structure of periodic stimuli may also be seen as a disadvantage. Aperiodic stimuli, by virtue of their randomly changing waveform, may have more idiosyncratic features to be remembered by, and therefore may enjoy an advantage in auditory memory. These hypotheses may be

---

Acknowledgment. This research was supported by NICHD Grant HD01994 and BRS Grant RR05596.

[HASKINS LABORATORIES: Status Report on Speech Research SR-69 (1982)]

called the engraving hypothesis and the redundancy hypothesis, respectively. Of course, both hypotheses could easily be wrong, and there may be no difference at all between periodic and aperiodic stimuli in memory. In that case, the differences between different stimulus sets in the studies cited above must have rested on stimulus properties other than periodicity.

We conducted a pilot study to compare stimuli that differ only in the presence versus absence of periodicity and are matched in all other respects. For that purpose, we generated band-limited complex nonspeech sounds by exciting the second-formant filter of a parallel-resonance speech synthesizer with either a periodic or an aperiodic source. The resulting stimuli were well matched in spectral shape and amplitude. The center frequency of the formant was varied in several steps, and the stimuli were presented in a fixed-standard same-different pitch discrimination task at two different interstimulus intervals. Differences between the two stimulus sets could emerge either as a difference in overall discrimination accuracy or as a difference in the effect of increasing the interstimulus interval (i.e., in the decay rate of auditory memory).

Method

Subjects. The subjects were paid volunteers recruited by advertisements on Yale campus. After a number of them had been tested, it became evident that several listeners made hardly any errors. These listeners (4 in all) were replaced with new subjects until a total of 12 had been run.

Stimuli. The stimuli were generated on the Haskins Laboratories parallel resonance synthesizer. Only the second-formant circuit was used. For the periodic stimuli, the "buzz" source was employed, and for the aperiodic stimuli, the "hiss" source. The fundamental frequency of the buzz was 100 Hz. Each set contained five stimuli with different center frequencies; nominally, they were 1611, 1688, 1764, 1840, and 1917 Hz. The nominal bandwidth was 90 Hz. To control for possible idiosyncrasies of the aperiodic stimuli due to the random noise source, three different tokens of each of the five aperiodic stimuli were synthesized. Amplitude parameters were set so as to yield equal amplitudes for periodic and aperiodic stimuli at output.

All stimuli were synthesized at a duration of 65 msec. Subsequently, they were digitized at 10 kHz using the Haskins Laboratories pulse code modulation system. The digitized waveforms were trimmed to a duration of 50 msec. This was done to eliminate artifacts at stimulus onset produced by starting the synthesizer at full amplitude. The periodic stimuli were reduced by eliminating the first pitch period and a portion from the end, so that exactly 5 complete, equal-amplitude pitch periods remained. The aperiodic stimuli were cut at corresponding points. To avoid transients, all cuts were made at the nearest zero-crossing.

Before recording the experimental tapes, the stimuli were analyzed using both standard spectrograms and spectral cross-sections generated by a Federal Scientific UA-6A spectrum analyzer. The spectral envelopes of periodic and aperiodic stimuli with the same center frequency were closely matched. We noted that the actual center frequencies did not always match the intended ones, but these discrepancies (which may have been due, in part, to inaccuracy of the synthesizer and, in part, to measurement error in spectral analysis) essentially left the spacing of the stimuli intact and were equally present in both types of stimuli.

Four tapes were prepared for the experiment. They contained pairs of either periodic or aperiodic stimuli at one of two interstimulus intervals (ISIs) between the members of each pair (.5 or 2 sec). The first member of each pair was constant; it was always the center stimulus of the continuum (1764 Hz). The second stimulus could be any of the five stimuli, with equal probability. Thus, on 20 percent of the trials, the two stimuli in a pair were identical; on 40 percent, the second stimulus had a higher pitch than the first, and on the remaining 40 percent, it had a lower pitch. There were two degrees of pitch difference, depending on whether the comparison stimulus was one or two steps removed from the standard. The five possible pairs occurred 30 times in random order, with 2 sec of silence between pairs for the short ISI and 4 sec for the long ISI. There were longer pauses after groups of 30. Different tokens of the aperiodic stimuli were employed in successive blocks of 5 pairs.

Procedure. The four stimulus tapes were presented in a single session. Their order was counterbalanced across subjects. The tapes were played back on an Ampex AG-500 tape recorder, and the subjects listened over TDH-39 earphones. The task was to write down "s" or "d" for each trial, depending on whether or not a difference could be detected between the two members of a pair. Each tape was preceded by several practice trials.

Analysis. Individual subject scores in the four different conditions were converted into d' values. The proportion of "same" responses to identical pairs was taken as the false-alarm rate, and the proportions of "different" responses to the four types of nonidentical pairs were taken as separate hit rates. The use of d' values corrected for differences in subjects' tendency to say "same"; however, the d' values were confounded with possible differences in subjects' criteria for detecting upward and downward changes in pitch. Proportions of 0 and 1 were treated as .01 and .99, respectively, leading to an upper bound on d' of 4.66.

A four-way analysis of variance was conducted on the d' values, with the factors Stimulus Type (periodic vs. aperiodic), ISI (.5 vs. 2 sec), Direction of Pitch Change (up vs. down), and Extent of Pitch Change (1 vs. 2 steps).

Results

The main results are shown in Figure 1a. There we see that discrimination performance declined as the ISI was extended from .5 to 2 sec, $F(1,11) = 5.6$, $p < .05$. However, there was no significant difference between the two types of stimuli, either in overall performance level or in the extent of the decline.

The effects of the other two factors, direction and extent of pitch change, are shown in Figure 1b. As expected, 2-step differences were easier to discriminate than 1-step differences, $F(1,11) = 61.5$, $p < .001$, but there was an interaction, $F(1,11) = 5.0$, $p < .05$: 1-step changes were easier to detect when pitch went down rather than up, whereas there was no effect of

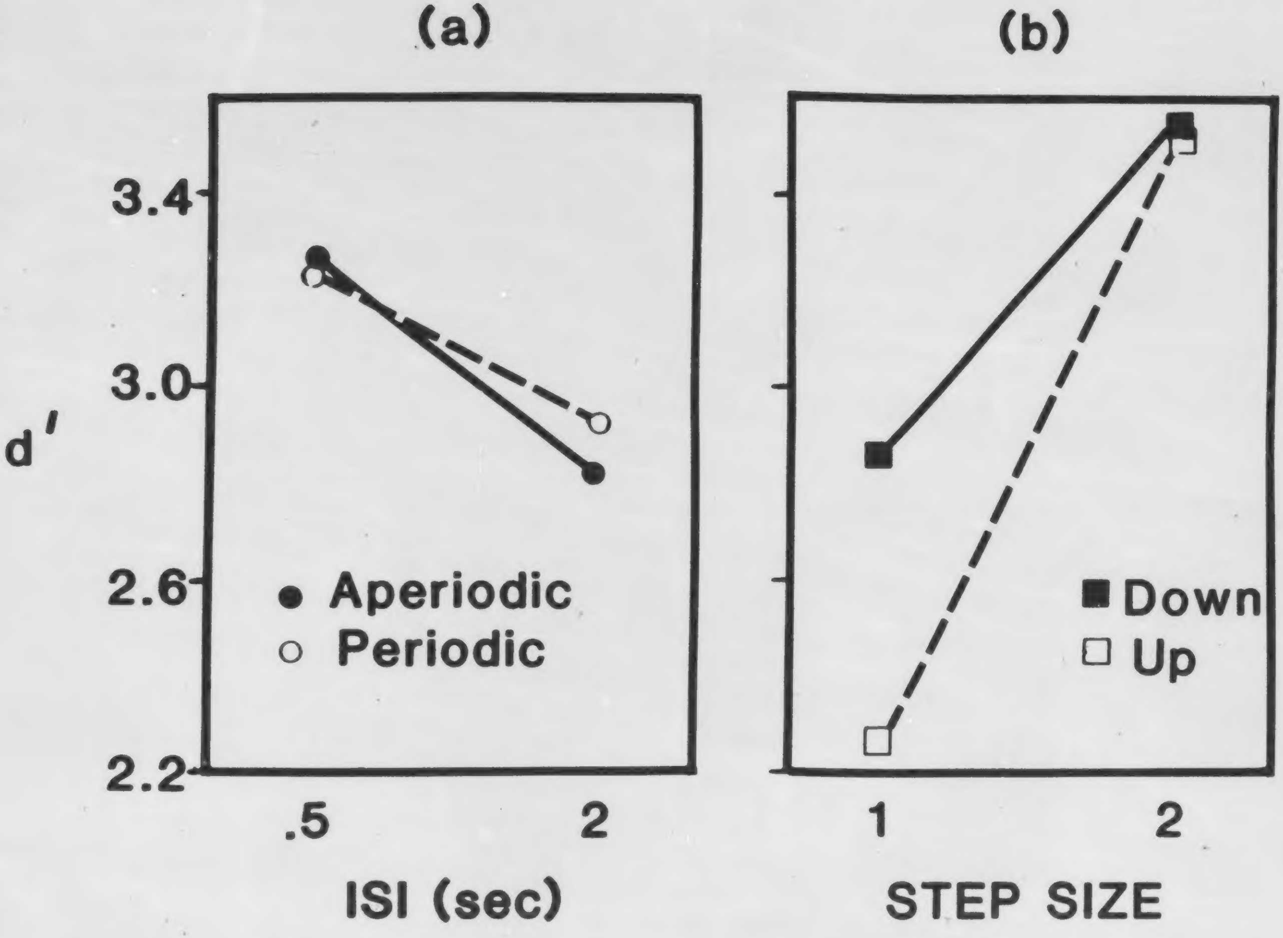


Figure 1. (a) Effect of ISI on memory for periodic and aperiodic stimuli. (b) Effects of magnitude (in steps) and direction of change on pitch discrimination.

direction in 2-step changes. This interaction may, in part, derive from a ceiling effect for 2-step pairs. The main effect of direction fell short of significance, $F(1,11) = 4.5$, $p > .05$. Note that better discrimination of downward changes would be expected on the basis of Weber's law.

One additional effect was significant in the analysis of variance. This was the three-way interaction between Stimulus Type, ISI, and Direction of Change, $F(1,11) = 5.0$, $p < .05$. The performance decline with ISI was larger for downward than for upward changes in aperiodic stimuli, but larger for upward than for downward changes in periodic stimuli. The reason for this pattern is not clear.

Discussion

Our results suggest that periodic and aperiodic stimuli are about equally well retained in auditory memory. However, we cannot prove the null hypothesis. It is possible that our experimental design did not provide the best opportunity for differences between the two types of stimuli to emerge. At least three reasons could be envisioned. First, the stimuli were relatively easy to discriminate--so much so that the best subjects had to be replaced. Thus, the data derive only from subjects with average or below-average discriminatory capabilities--an undesirable state of affairs. Second, the stimuli were rather brief, and it may be argued that their duration was too short for any specific advantage of periodic stimuli to arise. In other words, five pitch periods may not be enough to produce a sufficient amount of reinforcement-through-repetition, the process assumed by the "engraving hypothesis." Third, by using a fixed-standard paradigm, we may have reduced the subjects' reliance on auditory memory. In principle, subjects could have adopted the strategy of ignoring the standard altogether (i.e., by relying on a long-term memory representation of it) and of arriving at a "same-different" decision by an absolute judgment of the comparison stimulus. That this strategy was not used exclusively is suggested by the significant effect of increasing the ISI. Nevertheless, it may well be true that we did not force the listeners sufficiently to rely on auditory memory.

Thus, our pilot study did not put the engraving/redundancy hypotheses to a very strong test, and further research will be necessary to decide whether either of them has any validity. However, before further experiments are undertaken, we should perhaps wait for a more compelling reason to expect any effect of periodicity on auditory memory. The present results have not increased our confidence that such effects exist.

REFERENCES

Healy, A. F., & Repp, B. H. Context sensitivity and phonetic mediation in categorical perception. Journal of Experimental Psychology: Human Perception and Performance, 1982, 8, 68-80.

May, J. The perception of Egyptian Arabic fricatives. Unpublished doctoral dissertation, University of Connecticut, 1979.

READING SKILL AND LANGUAGE SKILL

Virginia A. Mann+

Abstract. To learn to read is to acquire a visual language skill that systematically maps onto extant spoken language skills. Some children perform this task quite adeptly, while others encounter much difficulty, and it has become a question of both scientific and practical merit to ask why there exists such a range of success in learning to read. Obviously, learning to read places a complex burden on many emerging capacities, and in principle, at least, reading disability could arise at any level from general cognition to visual perception. Yet since reading is parasitic on spoken language, the possibility also exists that reading disability is derived from some subtle difficulty in the language domain. In this article, my intent is to review some of the many studies that have explored the association between reading skill and spoken language skill. These reveal that when certain spoken-language skills of good and poor beginning readers are critically examined, considerably many, though perhaps not all, poor readers prove to possess subtle deficiencies that correlate with their problems in learning to read.

LINGUISTIC SHORT-TERM MEMORY DIFFERENTIATES GOOD AND POOR BEGINNING READERS

One of the more compelling reasons to view reading deficiency as the derivative of a language deficiency is that success at learning to read is associated with the adequacy of certain linguistic short-term memory skills. In our work at Haskins Laboratories, my colleagues and I have found clear indications of this association in a variety of different studies of good and poor beginning readers. For the moment, however, let me put aside a discussion of those studies in order to consider first the short-term storage requirements of normal language processing, and to summarize some recent findings as to how these requirements are met by the mature language user. These considerations pertain to both written and spoken language and provide a

---

+Also Bryn Mawr College.

Acknowledgment. Much of the work described in this paper, along with preparation of the manuscript itself, was supported by NICHD Grant HD-01994 and BRS Grant RR-05596 to Haskins Laboratories. I am grateful to the many members of the reading research group at Haskins, particularly to Isabelle Liberman and Donald Shankweiler for their comments on earlier drafts of this paper, and to Alvin Liberman and Ignatius Mattingly for their insights into the psychology of language.

[HASKINS LABORATORIES: Status Report on Speech Research SR-69 (1982)]

necessary introduction to any discussion of linguistic short-term memory among beginning readers.

An adequate short-term memory is essential to language comprehension simply because the component words of a phrase or sentence must often be held temporarily, pending extraction of the meaning of the whole phrase or sentence (Baddeley, 1978). It is for precisely this reason that many current models of sentence processing explicitly include some form of short-term memory buffer as a part of their parsing device (cf. Frazier & Fodor, 1978; Kimball, 1975; Marcus, 1980). Some consideration has been given to the form of memory representation that mediates human parsing. Current psychological theory has it that some level of phonetic representation is likely to be involved, this being an abstract representation of the articulatory gestures that constitute the material being parsed (Liberman, Mattingly, & Turvey, 1972). There are many experimental findings to corroborate this view. On the one hand, adult subjects have given evidence of relying on phonetic representation while performing such ecologically invalid tasks as recalling a string of letters or a string of words (Conrad, 1964; Drewnowski, 1980). More importantly, there is evidence that phonetic representation is also involved during comprehension of both written and spoken sentences (cf. Baddeley, 1978; Daneman & Carpenter, 1980; Kleiman, 1975; Levy, 1977; Slowiaczek & Clifton, 1980; Tzeng, Hung, & Wang, 1977).

It is, of course, not inconceivable that, in reading, some nonlinguistic representation of written words might be employed in lieu of a phonetic one (cf. Kleiman, 1975; Meyer, Schvaneveldt, & Ruddy, 1974). There is, after all, much evidence to suggest that access to the mental lexicon for printed words may not necessarily require reliance on phonetic representation (cf. Baron, 1973; Kleiman, 1975; Meyer et al., 1974). Nonetheless, it is important to emphasize that reading typically involves more than mere lexical access alone. A successful reader must often go beyond the lexicon and place reliance on the grammatical structure of the material being read. In contrast to experiments involving lexical access, those experiments concerned with reading situations where sentence structure is at stake have consistently given evidence of the involvement of phonetic representation (Daneman & Carpenter, 1980; Kleiman, 1975; Levy, 1977; Slowiaczek & Clifton, 1980). Even readers of Chinese logography, an orthography in which access to the lexicon is necessarily mediated by non-phonetic representation, appear to make use of phonetic representation when their task involves recovering the meaning of written sentences and not simply words alone (Tzeng et al., 1977).

For adult subjects, phonetic representation is clearly involved in both written and oral language comprehension. Having made this point, let me return to the primary concern of this paper, which is a review of some recent studies of good and poor beginning readers. These provide another form of support for the involvement of phonetic representation in all language processing, by revealing that effective use of phonetic representation is associated with, and may even presage success in, learning to read. I intend to review some of the many findings that support this conclusion; however, it might be useful first to provide some basic information about the population of beginning readers whom my colleagues and I have studied, since they have provided much of the data to which I will refer.

Most frequently our subjects have been first, second, and third graders who attend public schools. All of them are native speakers of English who suffer from no known neurological impairment. They are identified by their teachers as being "good," "average," or "poor" readers, a status that we confirm by administering standard reading tests to each child (typically the Word Attack and Word Recognition Subtests of the Woodcock Reading Mastery Tests, Woodcock, 1973; or the Word Recognition Subtests of the Wide Range Achievement Test, Jastak, Bijou, & Jastak, 1965). Administration of these tests has typically revealed the "good" readers to be reading at a level one or more years above their grade placement, whereas the "average" readers are reading at a level between one year above and one-half year below placement. The "poor" readers tend to be reading at a level one-half year or more below grade placement. Aside from administering standard reading tests, we have also usually given our subjects intelligence tests (either the Peabody Picture Vocabulary Test, Dunn, 1959; or the Slosson Intelligence Test for children, Slosson, 1963; or the WISC-R), and have excluded those children in either reading group who score below 90 or above 145.

One of the more general findings to emerge from our work is that good and poor readers may differ in temporary memory for some types of material, but not for other types (Katz, Shankweiler, & Liberman, in press; Liberman, Mann, Shankweiler, & Werfelman, in press; Mann & Liberman, in press). An example of this trend may be seen in the results of a study that assessed recognition memory skill among good and poor beginning readers (Liberman et al., in press). The subjects were second graders who differed in reading ability, but not in mean age or mean IQ. They participated in an experiment that employed the recurring recognition memory paradigm of Kimura (1963) as a means of evaluating memory for several different types of material. The material we studied included two non-linguistic materials--photographs of unfamiliar faces and nonsense "doodle" drawings--and one linguistic material--printed nonsense syllables. For each of these, the children inspected a set of stimuli and proceeded to indicate any of the inspection items that recurred in a subsequent recognition set. As may be seen in Figure 1, the poor readers were equivalent to the good readers in memory for faces and even somewhat better than the good readers (although not significantly so) in memory for the nonsense drawings. However, they were significantly inferior to good readers in memory for the nonsense syllables. Thus there is an interaction between reading ability and the type of item being remembered; an interaction that prevailed in an analysis of covariance adjusting for any effects of age or IQ differences.

Clearly, this experiment cannot support a conclusion that poor readers suffer from some general memory difficulty. Rather, they appear deficient only in the ability to remember linguistic material. Many findings that concern short-term memory lend further support to this conclusion. Good readers typically surpass poor readers in short-term memory for printed strings of letters or printed words (cf. Shankweiler, Liberman, Mark, Fowler, & Fischer, 1979; Mark, Shankweiler, Liberman, & Fowler, 1977) as well as for printed nonsense syllables. However, good readers also excel at recall of spoken strings of letters (Shankweiler et al., 1979), spoken strings of words (Bauer, 1977; Byrne & Shea, 1979; Katz & Deutsch, 1964; Mann, Liberman, & Shankweiler, 1980; Mann & Liberman, in press), and even spoken sentences (Mann et al., 1980; Perfetti & Goldman, 1976; Wiig & Roach, 1975; Weinstein &

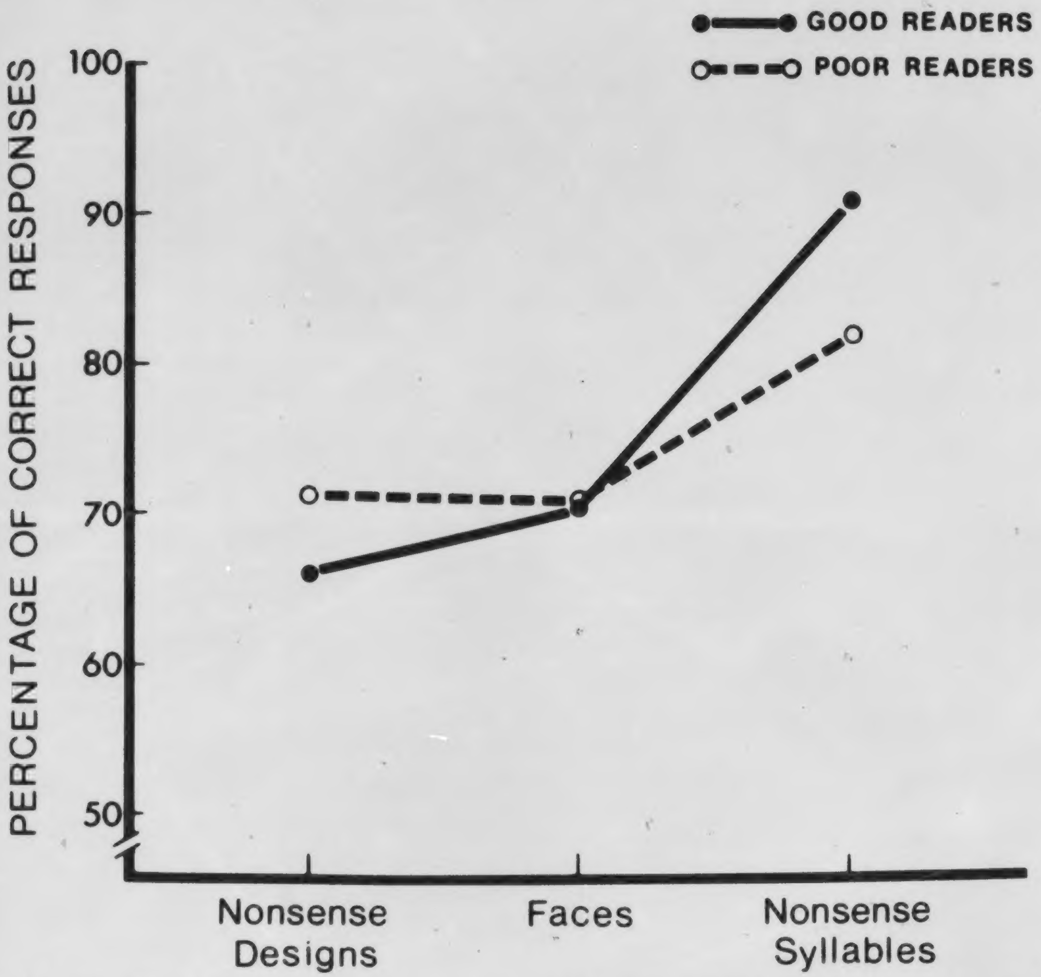


Figure 1. Good and poor readers' mean percentage of correct responses on nonsense designs, faces, and nonsense syllables.

Rabinovitch, 1971). At this point it is important to note that, since the advantage of good readers holds for both written and spoken material, it must extend beyond processes involved in reading, as such, to the broader realm of language processing.

To account for the linguistic memory distinctions between good and poor readers, some of my colleagues (Liberman & Shankweiler, 1979; Shankweiler et al., 1979) offered the hypothesis that poor readers have some difficulty that specifically compromises effective use of phonetic representation. Therefore, they used a modification of Conrad's (1964) procedure for examining the involvement of phonetic representation in memory for written letter strings, to test a group of good, average, and poor readers from a second-grade population that was homogeneous with respect to age and IQ. As was the case in Conrad's procedure, the children were asked to recall strings of five consonants that were of two basic types. Half of the strings were composed of consonants with phonetically confusable (i.e., rhyming) names, whereas the other half contained letters with phonetically nonconfusable (i.e., nonrhyming) names. During testing, the children saw a letter string with all of its letters printed in upper case on a single line in the center of the visual field. After a three-sec inspection period, when the letters could no longer be seen, they wrote down any letters that could be remembered, preserving the sequence as closely as possible.

On the basis of Conrad's findings, Liberman, Shankweiler, and their colleagues predicted that nonrhyming letter names would generate fewer phonetic confusions than rhyming ones, and thus facilitate recall in subjects who rely on phonetic representation as a means of retaining letters in short-term memory. It was felt that if a subject's level of performance failed to profit from reduced phonetic confusability, then that subject might have made less effective use of phonetic representation as a mnemonic device. The performance of good, average, and poor readers on the two types of letter strings is compared in the top section of Figure 2. Good readers, in general, made fewer errors than poor readers, and the average readers fell in between. The performance of the good readers, however, was also more significantly affected by the manipulation of rhyme than was that of the average or poor readers. In fact, the advantage of the superior readers was virtually eliminated when the letter strings contained letters with phonetically confusable names. In other words, phonetic confusability penalized the better readers to a greater extent than children in the other two reading groups.

These findings were extended by two subsequent experiments involving the same group of subjects and the same set of letter strings. In the first of these, the letters of each string were presented visually, but successively rather than simultaneously. In the second experiment, the letters were presented successively, but auditorily rather than visually. The results of these experiments are also displayed in Figure 2, where it may be seen that, once again, the interaction betwen reading ability and the effect of phonetic confusability was upheld. Indeed, it prevailed even when the letters were heard instead of seen. It is important to underscore the fact that reading ability was the only variable that interacted with the effect of phonetic confusability on letter recall. The children with higher IQ scores did tend to perform at a higher level than those with lower scores; however, the extent of their superiority was the same regardless of whether the comparison

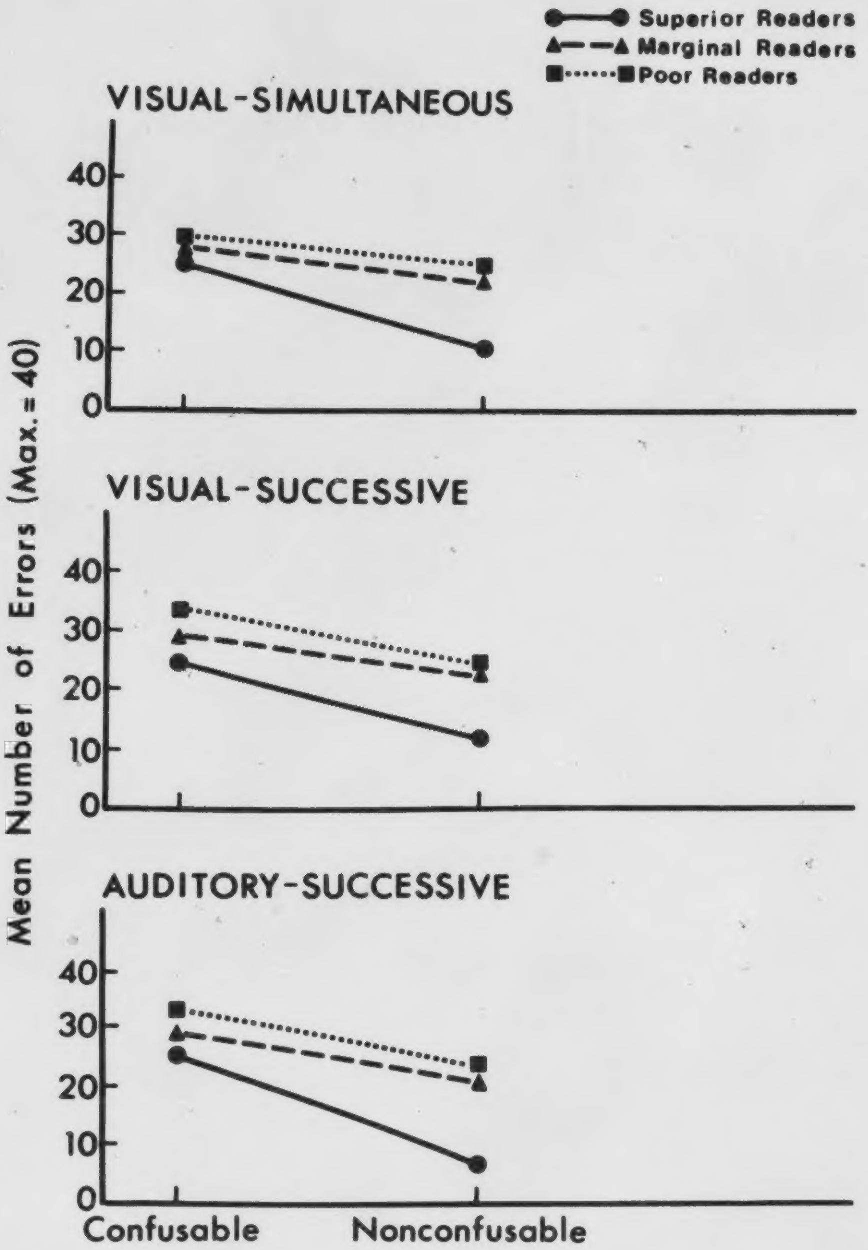


Figure 2. Mean errors of superior and good readers on recall of letter strings, summed over serial positions. (Means from delay and nondelay conditions are averaged. Maximum = 40.)

involved phonetically confusable letter strings or phonetically nonconfusable ones. Thus, the interaction between reading ability and the effect of phonetic confusability was unaltered when the analysis of the data covaried for any effects of IQ.

To strengthen these findings about poor readers' ineffective use of phonetic representation, my colleagues and I followed the study of letter-string recall with a study of the role of phonetic representation in recall of other, more ecologically valid material such as spoken word strings and spoken sentences (Mann et al., 1980). In that study, the subjects were again good and poor readers from a second-grade classroom. This time, however, the good readers had a slightly higher mean IQ than the poor readers. The experiment involved having the children in each group repeat strings of five spoken words, and also the words of 13-word sentences that were either meaningful or semantically anomalous. The materials included many different items of each type, but for word strings and both types of sentences, half of the items contained a high density of phonetically confusable (i.e., rhyming) words, whereas half contained phonetically nonconfusable words instead. Children's performance on the word strings is compared in Figure 3, and that on sentences is compared in Figure 4. As can be seen in those figures, for word strings, as well as for both meaningful and semantically anomalous sentences, good readers made fewer errors than poor readers as long as the material was phonetically nonconfusable. For all three types of material, however, they fell to the level of the poor readers when the material contained a high density of phonetically confusable words. In this experiment, although good readers tended to have higher IQ's, a significant interaction between reading ability and the effect of phonetic confusability was obtained when the results were subjected to an analysis of covariance that adjusted for any differences in IQ. Once again, intelligence alone was not the source of the good readers' more effective use of phonetic representation.

Thus, whether the material is apprehended by ear or by eye, and whether it involves letter strings or meaningful sentences, the performance of good readers tends to be both superior to that of poor readers and also more strongly affected by manipulations of phonetic confusability. For most good readers, as for most adults, phonetic confusability of the material to be recalled makes reliance on phonetic representation a liability rather than an asset. In contrast, phonetic confusability has little effect on the memory performance of most poor readers, a fact that we interpret as evidence that they are, for some reason, encountering difficulty with phonetic representation.

CLARIFYING THE BASIS OF POOR READERS' PROBLEMS WITH LINGUISTIC SHORT-TERM MEMORY

At this point, it becomes appropriate to consider why good and poor readers might differ in performance on tasks that involve reliance on phonetic representation. We can lay aside the possibility that memorial representation, in general, is a problem, since if this were so, poor readers would have been inferior on other tests of temporary memory and not merely on those that involve reliance on phonetic representation. A general cognitive deficiency would also seem an unlikely basis, given our findings that IQ scores are not

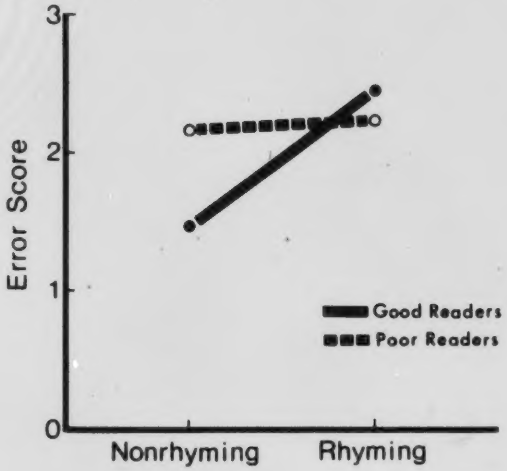


Figure 3. Mean error scores of good and poor readers on recall of word strings, in nonrhyming and rhyming conditions. (Maximum = 5.)

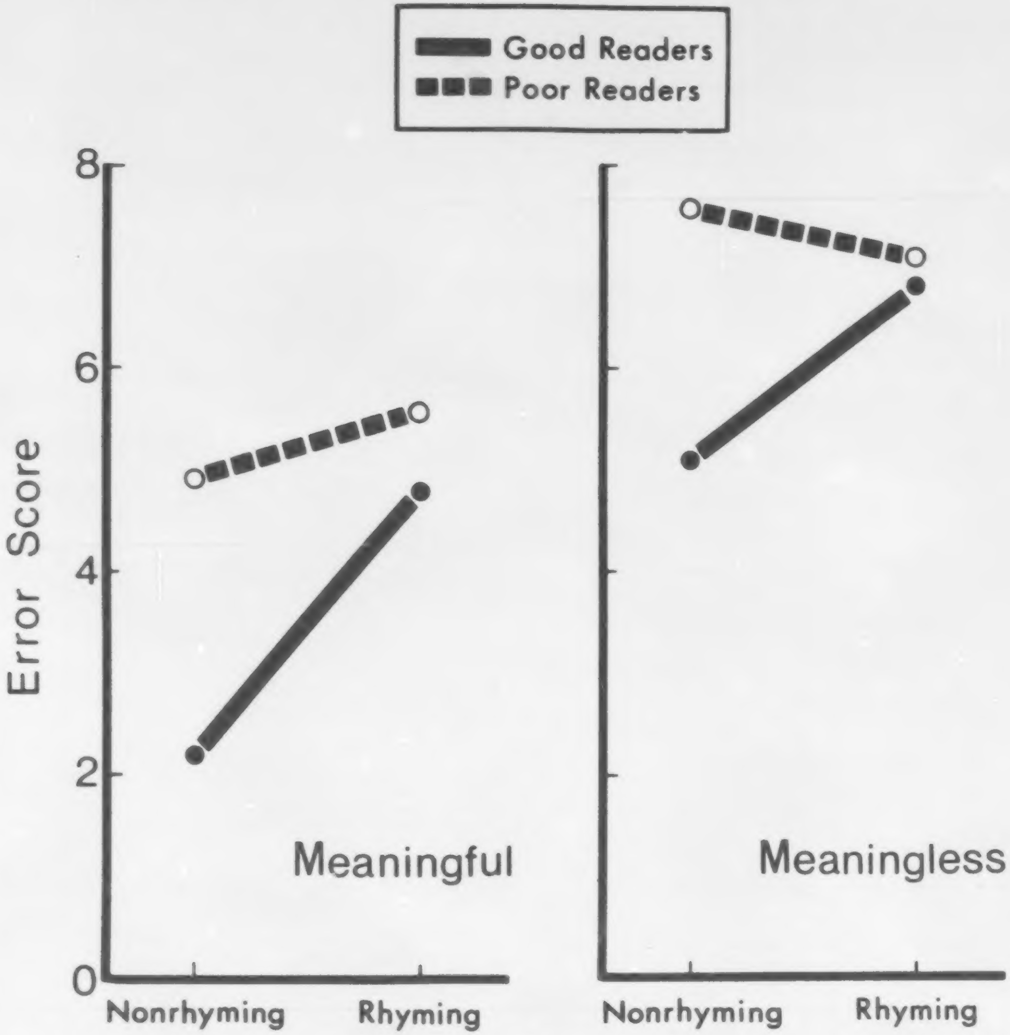


Figure 4. Mean error scores of good and poor readers on recall of meaningful and meaningless sentences in nonrhyming and rhyming conditions. (Maximum = 13.)

significantly associated with sensitivity to manipulations of phonetic confusability. Two other possibilities seem more plausible. On the one hand, poor readers might not resort to phonetic representation at all, relying instead on visual or semantic modes of representation. However, it is likewise possible that they do attempt to employ phonetic representation, but for some reason their representations are less effective.

One piece of evidence that is relevant to this issue is provided by the results of an experiment in which I extended Liberman and Shankweiler's study of letter string memory to a population of second- and third-grade children who were learning to read Dutch. The subjects were the ten best readers and the ten worst readers in each grade; their mean ages and reading abilities are given in Table 1. The procedure was the same as in the first experiment of Shankweiler et al. (1979) with one innovation. In constructing the letter strings, I separately manipulated phonetic and visual confusability, since this was more feasible in Dutch than in English. Thus it was possible to examine recall of three different types of upper-case consonant strings: strings of letters that were phonetically confusable but not visually confusable; strings of letters that were visually confusable but not phoneti-

Table 1

Age and Reading Ability Among Beginning Readers of Dutch

	Mean Age in Years	Mean Grade-equivalent Reading Ability*
Second Graders:		
good readers	7.4	4.1
poor readers	7.3	1.1
Third Graders:		
good readers	8.2	5.5
poor readers	8.5	2.3

*Grade-equivalent scores measured by the Ein-Minuten Test; Berkhout, 1972.

cally confusable, and strings of letters that were minimally confusable along both the visual and phonetic dimension. In all cases, the measure of phonetic confusability was the density of letters with rhyming names, since that measure had been employed by the Conrad (1964) study on which the Shankweiler et al. (1979) study had been based. The measure of visual confusability was derived from the upper-case letter confusion matrix compiled by Townsend (1971), and was the summed probability of visual confusion for each possible pair of letters in a given string. Computed in this way, the mean confusability for the ten visually confusable strings was 0.81, and was significantly greater than that for either the ten phonetically confusable or the ten minimally confusable strings (0.27 and 0.31, respectively, $t(18)=3.1$, $p<.01$, and $t(18)=2.8$, $p<.01$, respectively).

As no children's IQ test was available in Dutch, I controlled for nonlinguistic short-term memory rather than for general intellectual ability. The test of nonlinguistic memory that I administered was the Corsi test (Corsi, 1972). The materials for that test consist of a set of nine wooden cubes attached in a random fashion to a flat wooden base. The entire apparatus is painted black; there are identifying numbers on the rear surface of the cubes that can be seen by the experimenter although not by the subject. During testing, the subject watches the examiner tap out a sequence of blocks and then attempts to reproduce that sequence. Practice sequences of two and three blocks are given first, followed by eight test sequences of four and eight of five blocks each. The suitability of this test as a measure of nonlinguistic short-term memory is indicated by clinical studies revealing that whereas performance on linguistic short-term memory tests is selectively impaired by damage to the left or language-dominant hemisphere, that on the Corsi blocks shows the opposite pattern of selective impairment as a consequence of damage to the right, or language-nondominant hemisphere (Corsi, 1972; Milner, 1972).

Because of my experience with American children, which had revealed no significant relation between reading ability and non-linguistic memory, I did not anticipate finding that good and poor beginning readers of Dutch would differ in performance on the Corsi test. There seemed to be no reason to anticipate that children in the two reading groups would differ in nonlinguistic abilities. It did seem possible, however, that poor readers would do less well than good readers on the letter-string memory test, and that they might also be differently affected by the manipulations of phonetic and visual confusability. Proceeding from the fact that phonetic confusability penalizes recall in subjects who rely on phonetic representation, I speculated that if poor readers rely on visual representation, then they might be inordinately affected by the manipulation of visual confusability.

The results of the study are given in Table 2, where all memory test scores are error scores that include errors of item omission and substitution, as well as of incorrect order. In that table, it may be seen that despite any differences in the Dutch and English languages or in the educational practices by which they are taught, the memory profiles of good and poor readers in the two countries prove quite similar. As we have found to be the case for American children, Dutch children who are poor readers are equivalent to good readers in performance on the nonlinguistic short-term memory test: $F(1,39)=1.6$, $p>.10$, although older children tend to do better than younger ones: F(1,39)=4.9, p<.05. Older children also tended to do better on the letter string test: F(1,39)=11.8, p<.005. More importantly, the good beginning readers of Dutch tended to surpass the poor readers in memory for consonant strings, and they did so at both age levels tested: F(1,39)=45.0, p<.0001.

Table 2

Results Obtained with Beginning Readers of Dutch

	CORSI BLOCKS	LETTER STRINGS Phonetically Confusable	LETTER STRINGS Visually Confusable	LETTER STRINGS Non-confusable
Second grade:				
good readers	20.5	15.4	23.4	16.4
poor readers	24.5	29.8	29.8	30.9
Third grade:				
good readers	15.9	11.0	20.7	8.5
poor readers	18.2	24.5	26.9	26.2
	(Max. = 72)	(Max. = 50)	(Max. = 50)	(Max. = 50)

In these data, there is, further, the anticipated interaction between reading ability and the effect of the various manipulations of letter confusability: F(2,72)=28.3, p<.0001. The better readers surpassed the poorer ones in memory for the minimally confusable letter strings, this being true for both second: t(18)=4.8, p<.001, and third graders: t(18)=10.3, p<.001. However, the good readers at both ages fell to the level of poor readers when they attempted to recall phonetically confusable strings. A further twist to these data involves the effect of visual confusability, or rather, its non-effect. Neither good nor poor readers were affected by the presence of a higher density of visual confusability. That is to say, for both groups of subjects at both age levels, performance on the visually confusable strings was no different from that on the nonconfusable ones. This gives us no reason to believe that in this task the poor readers opted for a purely visual representation of the letter strings. Either they relied on some as yet undetermined form of representation, or they relied on phonetic representation and for some reason failed to profit from reduced phonetic similarity among the letter names.

Some direct evidence in support of the possibility that poor readers do sometimes rely on phonetic representation may be found in the pattern of errors these children make when they attempt to recall a phonetically confusable string of spoken words. Some of my colleagues and I recently analyzed the responses made by good and poor readers who were attempting to recall such a string (Brady, Shankweiler, & Mann, 1982). The subjects were participating in an experiment that will be described in more detail below; they were good and poor readers from a third-grade classroom and they did not significantly differ in IQ. They were asked to repeat strings of five words that were either phonetically confusable or phonetically nonconfusable. As in the past, the good readers tended to excel with respect to the poor readers, but also tended to be more greatly affected by the manipulation of phonetic confusability. We also found, however, that although children in both reading groups made many substitution errors, the poor readers tended to make more of these than the good readers. We therefore turned to analyzing the composition of the substitution errors and their relation to the words of the original string.

Our analysis revealed that the pattern of substitution errors was the same for good and poor readers alike. Almost no substitutions were semantic associates of the words in the string being recalled; instead, the majority were composed of a subset of the phonemes that had constituted the words of the string being remembered. For example, a great proportion of the errors contained an appropriate initial consonant and even more contained an appropriate vowel or final consonant. Thus it seemed as if the children in both reading groups had remembered many of the phonemes they had heard. The poor readers, for some reason, had merely made more errors in recalling the original word strings, perhaps because their phonetic representations were less well formed, or perhaps because their representations decayed more rapidly than those of the good readers.

Thus, in at least some circumstances, it seems that poor readers may rely on phonetic representation to some extent; otherwise they would not have tended to make substitution errors that preserve phonetic aspects of the original word string. Before leaving this topic, it would be pertinent to mention the possibility that problems with phonetic representation may force the poor readers to rely on semantic representation during certain memory tasks. Although my colleagues and I have seen almost no semantically-based substitution errors among either good or poor readers, this has not been the case in another study done by Byrne and Shea (1979). These investigators compared the performance of good and poor beginning readers on a spoken-word recognition memory test, and found that, in general, good readers performed at a higher level than poor readers. They also discovered that children in the two groups tended to make different types of errors. Whereas poor readers made proportionately more false recognition errors on semantic associates of the correct items, good readers tended to make more such errors on words that were phonetic associates. For example, when asked to remember and subsequently recognize "home," poor readers tended erroneously to recognize "house," but good readers, "comb." Yet when the task was to remember nonsense syllables instead of words, children in both reading groups made many errors on phonetic foils. Once again, however, good readers somehow made more effective use of phonetic representation, as evidenced by their tendency to make fewer errors, in general, coupled with their tendency to make disproportionately many errors on phonetically-similar foils.

Turning now to the question of why the phonetic representations of poor readers may be less effective than those of good readers, let me return to the above-mentioned study by Brady et al. (1982). In that study an approach to the problem of phonetic representation was inspired by the finding that, when speech perception is stressed by the presence of background noise, short-term memory span is inordinately affected (Rabbitt, 1968). This finding led us to consider the possibility that the short-term memory difficulties of poor readers might be associated with some difficulties in encoding speech. Therefore, we designed an experiment to compare the ability of good and poor readers to identify spoken words that were partially masked by white noise.

The third graders who were subjects of this study did not differ in age or IQ, but did differ in reading ability, and also in memory for strings of spoken words. Their performance showed the usual interaction between reading ability and the effect of phonetic confusability. They were asked to identify a pre-recorded set of spoken words that contained an equal number of high and low frequency words and was balanced for phonetic constituents and syllabic structure. Each child heard the words under two different conditions: first partially masked by signal-correlated white noise, and later under more optimal listening conditions.

The results revealed that although the poor readers were not significantly different from good readers in performance under the optimal conditions, they made about 35% more errors when the words were partially masked. That this problem could not be attributed to some basic vocabulary deficiency could be seen from the fact that differences between children in the two reading groups obtained equally for high and low frequency words, and also from the fact that the subjects of our study had performed at the same level on the Peabody Picture Vocabulary Test (Dunn, 1959). It is also consistent with this observation that an interaction between reading ability and the effect of partial-masking was obtained with an analysis that covaried for the effects of age and IQ.

To determine whether the findings of this experiment were specific to speech perception, as opposed to being an attribute of general auditory perception, we conducted a second experiment. In it, the same subjects were asked to identify a set of environmental sounds taken from a standard clinical test, including such sounds as a cat meowing and a door slamming. The procedure was analagous to that in the previous experiment with spoken words; the subjects first identified the sound when partially masked by white noise, and later when presented under more optimal listening conditions. The pattern of results for this second experiment proved distinct from that obtained in the first one. Many of the poor readers were actually better than the good readers at identifying the partially-masked sounds, although this difference is not significant. An analysis of covariance that adjusted for age and IQ effects reveals that, although the noise penalized the overall level of performance, there was neither an effect of reading ability nor an interaction between reading ability and the penalizing effects of the noise masking.

Thus it would appear that any deficiency in auditory perception on the part of the poor readers is limited to the realm of speech perception. Although more research is needed to clarify the relation between this speech perception deficiency and poor readers' problems with phonetic representation, the fact of its existence is certainly provocative and most pertinent to the view that reading skill is associated with language skill.

LINGUISTIC SHORT-TERM MEMORY SKILL MAY PRESAGE READING SUCCESS

Having made a link between reading skill and effective use of phonetic representation in linguistic short-term memory tasks, and having reviewed some of the evidence as to why poor readers may have difficulty with phonetic representation, I will now concentrate on some ramifications of this difficulty. According to the view introduced in the beginning sections of this paper, phonetic representation is crucially involved in all normal language processing. Since spoken language antedates written language, and insofar as phonetic representation is involved in spoken language processing, difficulty with phonetic representation should often be found as an antecedent of reading failure.

A study completed only a short time ago speaks to this point, revealing that those kindergarten-aged children who make less effective use of phonetic representation in a word-string recall task are likely to become the poorer readers of their first-grade classrooms (Mann & Liberman, in press). The subjects for that study were a population of kindergarteners whom we followed longitudinally for one year. During May of the kindergarten year we assessed their memory for spoken strings of phonetically confusable and nonconfusable words, their memory for nonlinguistic material (the Corsi block sequences), and their awareness of the syllabic structure of spoken words. The following year, as first graders, these same children again received all of the memory tests, and a standard reading test. At this time they were rated by their teachers as "good," "average," or "poor" in reading ability.

The findings for the two years of the study are given in Table 3. Note first that the children in the three reading groups had equivalent IQ scores; we found no correlation between IQ scores and our measures of reading achievement. The children in the three groups also performed equivalently on the Corsi test of nonlinguistic memory; neither their kindergarten nor their first grade scores on this test were correlated with our reading measure. In contrast, however, both of our linguistic measures proved able to distinguish between children in the three different reading groups. Elsewhere we have discussed the relation between success at learning to read and the ability to realize the syllabic structure of spoken words (see, for example, Liberman & Mann, in press; or Mann & Liberman, in press). Here I will focus on the relation between effective use of phonetic coding and reading skill. It can be seen in Table 3 that children in the three reading groups were strongly and significantly differentiated by their performance on the phonetically nonconfusable word strings. As first graders, children's performance on this type of word string was significantly correlated with their reading ability--more importantly, a significant correlation also existed between their kindergarten performance on the phonetically nonconfusable word strings, and their first-grade reading ability. Note further that both as kindergarteners and as first graders, the poorer readers tended not only to perform at the lower levels on the word string memory test, but also to be among those least affected by the manipulation of phonetic confusability. Thus, their ineffective use of phonetic representation not only associated with their difficulty in learning to read, but actually presaged it.

TABLE 3

Mean error scores of good, average and poor readers on memory tasks: A longitudinal study (IQ determined in kindergarten, reading achievement in first grade).

READING ABILITY	GRADE LEVEL	VERBAL MEMORY Max =32		NONVERBAL MEMORY Max = 32	SYLLABLE SEGMENTATION TASK
		Nonrhyming Word Strings	Rhyming Word Strings	Corsi Blocks	(Percent passed in Kdgn.)
GOOD READERS N = 26 IQ 114.7	KDGN	8.1	13.4	8.4	85%
	1st GRADE	5.5	12.1	8.7	
AVERAGE READERS N = 19 IQ 114.7	KDGN	12.8	15.4	9.0	56%
	1st GRADE	9.2	11.3	8.1	
POOR READERS N = 17 IQ 115.5	KDGN	13.2	15.0	10.1	24%
	1st GRADE	13.7	12.7	10.1	

READING SKILL, LINGUISTIC SHORT-TERM MEMORY AND ORAL COMPREHENSION

The finding that effective use of phonetic representation can be a precursor of reading success is consistent with the view that reading skill derives from language skill, given the position that effective language comprehension is linked to effective phonetic representation, and the presumption that successful comprehension is essential to learning to read well. Clearly, one final demonstration is called for. If poor readers tend to make less effective use of phonetic representation than good readers, and consequently encounter difficulty retaining the words of sentences, then we may be able to demonstrate that they are less able to comprehend spoken sentences, especially if comprehension demands reliance on an effective short-term memory store.

Together with some of my colleagues (Donald Shankweiler and Suzanne Smith) I am currently analyzing the results of a study that asks whether there exists a three-way link between reading skill, effective use of phonetic representation, and spoken language comprehension. Clearly the existence of such an association would further support the view that reading skill is a product of language skill. The subjects of this most recent study are good and poor readers from a third grade population that is homogeneous with respect to age and IQ. They have been given a test of memory for strings of phonetically nonconfusable words, and several different tests of oral language comprehension, including two tests of our own design and one standardized clinical test. Thus far, we have only completed our analysis of the results of the standardized tests, a test called the Token Test (DeRenzi & Vignolo, 1962).

In the Token Test, subjects receive a series of oral instructions that specify how they are to manipulate a set of small colored "tokens." It has enjoyed considerable success as a reliable indicator of disorders of oral comprehension both among patients with acquired language deficits (DeRenzi & Vignolo, 1962) and children with developmental language disorders (LaPointe, 1976). We chose to use it because it forces reliance on the grammatical structure of a sentence rather than on common-sense knowledge or extralinguistic cues, and also because it poses an obvious stress on short-term memory.

The test itself consists of five basic parts that are graded in complexity. For the first four parts, all of the instructions are simple imperative sentences that contain a constant verb and either one or two noun-phrase objects. The instructions systematically increase from part to part in the number of objects involved and in the adjectival content (one or two adjectives) of the noun phrase. For the fifth part, the instructions contain as many words or more than those in the third and fourth parts, but further contain a series of different verbs and different noun phrase structures in the predicate. Thus the first four parts of the test involve a systematic increase in the number of objects and attributes that the subject must remember, whereas the fifth involves not only a substantial memory load but also an increase in syntactic complexity.

In general, the results of our study of Token Test performance have revealed that poor readers tend to do less well than good readers. In particular, we find that they do as well as good readers on the first three parts of the test, but fall behind on the last two parts. We had anticipated that the fourth and fifth parts might pose relatively more difficulty for the poor readers, simply because they contain the longest instructions. However, we recognize that difficulty on the fifth part of the test could also be a consequence of a more specific difficulty with recovering syntactic structure, aside from a short-term memory deficiency. Thus, while we have indeed established a relation between reading ability and oral comprehension of sentences, it remains to be determined whether ineffective use of phonetic representation can account for this relation in any direct way. We have some indication that for the children whom we tested, performance on the Token Test was at least moderately correlated with word-string memory performance. It also appears possible that for both the good and poor readers, the errors made on part five may have been direct consequences of the memory demands posed by certain instructions. We hope to continue to gain more insight into this issue as we analyze the results of our other two comprehension tests.

As we pursue this and other research, my colleagues and I are entertaining several possible outcomes. On the one hand, ineffective phonetic representation could not only compromise ongoing sentence processing, but also limit the development of linguistic competence. It is also within the realm of possibility that poor readers possess a comprehension deficit that is not so much a consequence as a concomitant of difficulty with phonetic representation. Perhaps reading disability, ineffective phonetic representation, and comprehension deficiencies are all manifestations of some more general language impairment that we have only begun to characterize. Surely the characterization of that impairment will be a productive research objective, since it may both illuminate our understanding of the psychology of reading, and clarify our approach to the current epidemic of reading failure.

REFERENCES

Baddeley, A. D. The trouble with levels: A reexamination of Craik and Lockhart's framework for memory research. Psychological Review, 1978, 85, 139-152.

Baron, J. Phonemic staging not necessary for reading. Quarterly Journal of Experimental Psychology, 1973, 25, 241-246.

Bauer, R. Memory processes in children with learning disabilities. Journal of Experimental Child Psychology, 1977, 24, 415-430.

Brady, S., Shankweiler, P., & Mann, V. Speech perception and memory coding in relation to reading ability. Haskins Laboratories Status Report on Speech Research, 1982, SR-69, this volume.

Byrne, B., & Shea, P. Semantic and phonetic memory codes in beginning readers. Memory & Cognition, 1979, 7, 333-338.

Conrad, R. Acoustic confusions in immediate memory. British Journal of Psychology, 1964, 55, 75-84.

Corsi, P. M. Human memory and the medial temporal region of the brain. Unpublished doctoral dissertation, McGill University, 1972.

Daneman, M., & Carpenter, P. A. Individual differences in working memory. Journal of Verbal Learning & Verbal Behavior, 1980, 19, 450-466.

DeRenzi, E., & Faglioni, P. Normative data and screening power of a shortened version of the Token Test. Cortex, 1978, 14, 41-49.

DeRenzi, E., & Vignolo, L. A. The Token Test: A sensitive test to detect receptive disturbances in aphasia. Brain, 1962, 85, 665-678.

Drewnowksi, A. Memory functions for vowels and consonants: A reinterpretation of acoustic similarity effects. Journal of Verbal Learning and Verbal Behavior, 1980, 19, 176-193.

Dunn, L. L. Peabody picture vocabulary test. Circle Pines, Minn.: American Guidance Services, 1959.

Frazier, L., & Fodor, J. D. The sausage machine: A new two-stage parsing model. Cognition, 1978, 6, 291-325.

Jastak, J., Bijou, S. W., & Jastak, S. R. Wide range achievement test. Wilmington, Del.: Guidance Associates, 1965.

Katz, P. A., & Deutsch, M. Modality of stimulus presentation in serial learning for retarded and normal readers. Perceptual and Motor Skills, 1964, 19, 627-633.

Katz, R., Shankweiler, D., & Liberman, I. Y. Memory for item order and phonetic coding in the beginning reader. Journal of Experimental Psychology, in press.

Kimball, J. Predictive analysis and over-the-top parsing. In J. Kimball (Ed.), Syntax and semantics (Vol. 4). New York: Academic Press, 1975.

Kimura, D. Right temporal lobe damage. Archives of Neurology, 1963, 8, 264-271.

Kleiman, G. M. Speech recoding in reading. Journal of Verbal Learning and Verbal Behavior, 1975, 14, 323-339.

LaPointe, C. M. Token test performances by learning-disabled and achieving adolescents. British Journal of Disorders of Communication, 1976, 11, 121-133.

Levy, B. A. Reading: Speech and meaning processes. Journal of Verbal Learning and Verbal Behavior, 1977, 16, 623-638.

Liberman, A. M., Mattingly, I. G., & Turvey, M. Language codes and memory codes. In A. W. Melton & E. Martin (Eds.), Coding processes and human memory. Washington, D.C.: Winston and Sons, 1972.

Liberman, I. Y., & Mann, V. A. Should reading remediation vary with the sex of the child? Bulletin of the Orton Society, in press.

Liberman, I. Y., Mann, V. A., Shankweiler, D., & Werfelman, M. Children's memory for recurring linguistic and non-linguistic material on relation ability. Cortex, in press.

Liberman, I. Y., & Shankweiler, D. Speech, the alphabet and teaching to read. In L. B. Resnik & P. A. Weaver (Eds.), Theory and practice of early reading (Vol. 2). Hillsdale, N.J.: Lawrence Erlbaum Associates, 1979, 109-134.

Liberman, I. Y., Shankweiler, D., Liberman, A. M., Fowler, C., & Fischer, F. W. Phonetic segmentation and recoding in the beginning reader. In A. S. Reber & D. L. Scarborough (Eds.), Toward a psychology of reading: The proceedings of the CUNY Conferences. Hillsdale, N.J.: Erlbaum, 1977, 207-225.

Mann, V. A., & Liberman, I. Y. Phonological awareness and verbal short-term memory: Can they presage early reading problems. To appear in F. H. Duffy & N. Geschwind (Eds.), *Dyslexia: Current status and future directions*.

Mann, V. A., Liberman, I. Y., & Shankweiler, D. Children's memory for sentences and word strings in relation to reading ability. *Memory & Cognition*, 1980, *8*, 329-335.

Marcus, M. P. *A theory of syntactic recognition for natural language*. Cambridge, Mass.: M.I.T. Press, 1980.

Mark, L. S., Shankweiler, D., Liberman, I. Y., & Fowler, C. A. Phonetic recoding and reading difficulty in beginning readers. *Memory & Cognition*, 1977, *5*, 623-629.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. Functions of graphemic and phonemic codes in visual word-recognition. *Memory & Cognition*, 1974, *2*, 309-321.

Milner, B. Disorders of learning and memory after temporal-lobe lesions in man. *Clinical Neurosurgery*, 1972, *19*, 421-446.

Morrison, F. J., Giordani, B., & Nagy, J. Reading disability: An information processing analysis. *Science*, 1977, *196*, 77-79.

Perfetti, C. A., & Goldman, S. Discourse memory and reading comprehension skill. *Journal of Verbal Learning and Verbal Behavior*, 1976, *14*, 33-42.

Rabbitt, P. M. A. Channel-capacity, intelligibility and immediate memory. *Quarterly Journal of Experimental Psychology*, 1968, *20*, 241-248.

Shankweiler, D., Liberman, I. Y., Mark, L. S., Fowler, C. A., & Fischer, F. W. The speech code and learning to read. *Journal of Experimental Psychology: Human Learning and Memory*, 1979, *5*, 531-545.

Slosson, R. L. *Slosson Intelligence Test for children and adults*. East Aurora, N.Y.: Slosson Educational Publications, 1963.

Slowiaczek, M. C., & Clifton, C. Subvocalization and reading for meaning. *Journal of Verbal Learning and Verbal Behavior*, 1980, *19*, 573-582.

Townsend, J. T. Theoretical analysis of an alphabet confusion matrix. *Perception & Psychophysics*, 1971, *9*, 40-50.

Tzeng, O. J. L., Hung, D. L., & Wang, W. S-Y. Speech recoding in reading Chinese characters. *Journal of Experimental Psychology: Human Learning and Memory*, 1977, *3*, 621-630.

Weinstein, R., & Rabinovitch, M. S. Sentence structure and retention in good and poor readers. *Journal of Educational Psychology*, 1971, *62*, 25-30.

Wiig, E. H., & Roach, M. A. Immediate recall of semantically varied sentences by learning disabled readers. *Perceptual and Motor Skills*, 1975, *40*, 119-125.

Woodcock, R. W. *Woodcock reading mastery tests*. Circle Pines, Minn.: American Guidance Services, 1973.

ON THE ROLE OF SIGN ORDER AND MORPHOLOGICAL STRUCTURE IN MEMORY FOR AMERICAN SIGN LANGUAGE SENTENCES

Vicki L. Hanson and Ursula Bellugi+

Abstract. Sentence processing in a visual/gestural language was investigated by testing signers' recognition for American Sign Language (ASL) sentences. Using a continuous recognition paradigm, sign order and structural changes that altered the meaning of a sentence were noticed at both immediate and delayed (45 seconds) test intervals. Sign order and structural changes that resulted in a paraphrase of an earlier-occurring sentence were noticed only with immediate testing. These results indicate that signers decompose a complex sign into its lexical and inflectional components during sentence comprehension and remember the meaning expressed by these components rather than remembering the exact sign structure.

In the past decade, it has become clear that there are primary gestural systems, passed down from one generation of deaf people to the next, that have taken their own course of development as autonomous languages. American Sign Language (ASL) is the common form of communication used by deaf native signers among themselves across the United States and parts of Canada. ASL is a primary visual-gestural system, not based on, nor derived from, any form of English, having its own lexicon and grammar (Klima & Bellugi, 1979). Because of the radical difference in the transmission medium for signed languages as opposed to spoken languages, ASL affords an opportunity to examine a question not easily investigated in other ways: namely, to what extent is language processing shaped by the production modality?

Research on American Sign Language shows that this visual-gestural system exhibits formal structuring at the same two levels as spoken languages: a sublexical level of structure internal to the sign (the phonological level in spoken languages) and a level of structure that specifies the ways signs are

---

+The Salk Institute for Biological Studies.

Acknowledgment. We wish to thank Malinda Williams, Ben Bahan, and Sandra Fields for their help on this study. Appreciation is also extended to Nancy Frishberg, Horace Reynolds, and Dennis Schemenauer who made arrangements for facilities and participants. The cooperation of the following institutions made this research possible: Gallaudet College, New York University, and California State University, Northridge. Illustrations were drawn by Frank A. Paul. Helpful comments on an earlier draft of this manuscript were provided by Deborah Wilkenfeld. This research was conducted while the first author was a postdoctoral fellow at The Salk Institute. The research and writing of this paper were supported in part by National Institute of Health Research Service Award NS-06109, National Science Foundation Grant BNS 79-16423, and by NICHD Grant HD-01994.

bound together into sentences (the grammatical level). ASL thus shares underlying principles of organization with spoken languages, but the formal devices that appear arise from the very different possibilities afforded by the visual-gestural modality. In spoken languages, the structuring of lexical items and morphological processes is essentially sequential. Words are composed of sequentially produced sounds; morphological processes commonly involve affixation of phonemic segments sequentially ordered in the sound stream. In ASL, however, signs are composed of contrasting formational parameters, co-occurring throughout the sign, and the morphological processes of ASL involve embedding the sign stem in superimposed contours of movements. Sign and inflectional marker thus co-occur in time (Bellugi, 1980).

In relation to the structure of the signs themselves, it has been determined that signs are not just holistic and iconic, but rather are composed of a limited number of arbitrary formational components that combine in regular and constrained ways in the signs of the language (Battison, 1974; Stokoe, Casterline, & Croneberg, 1965). These parameters (handshape, location, and movement) have been found not only to be formal linguistic descriptions of signs, but also to be psychologically real in the sense that signers rely on these parameters in sign processing. In studies of short-term memory for signs, signers characteristically make errors based on these formational parameters (Bellugi, Klima, & Siple, 1975; Frumkin & Anisfeld, 1977). Similarly, "slips of the hand" involve exchanges of these components (Newkirk, Klima, Pedersen, & Bellugi, 1980).

ASL differs dramatically from spoken languages in the form of its morphological processes. As a visual-gestural language, its morphological devices embed sign stems in superimposed changes of space and movement. Figure 1, for example, shows the sign PREACH under a variety of morphological operations [e.g., PREACH (basic sign), 'preach to them,' 'preach to each of them,' 'preach regularly,' etc.].[1]

The wide variety of semantic distinctions that are obligatorily marked morphologically in ASL sentences are often indicated in English either lexically or phrasally. ASL verb signs, for example, undergo obligatory inflections for referential indexing, indicating subject and/or object of the verb; for reciprocity; for grammatical number, marking distinctions such as dual and multiple; for temporal aspect, indicating distinctions such as 'regularly,' 'over and over again,' 'for a long time,' 'gradually'; for distributional aspect 'to each,' 'to any,' 'all over,' 'to certain ones.' There are also a large number of derivational processes such as those that derive nouns from verbs, that derive predicates from nouns, and that signal figurative or extended meanings. The elaborate system of formal inflectional devices, their widespread use to vary the form of signs, and the variety of fine distinctions they systematically convey suggest that ASL, like Latin and Navajo, is one of the inflective languages of the world (Klima & Bellugi, 1979).

The present research examines for the first time whether these formal linguistic descriptions of ASL morphological structure correspond to psychological representations that signers use in the interpretation and retention of sentences.

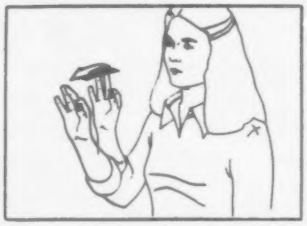
Several experiments with English have been conducted to test whether linguistic descriptions of the language can be applied to describe the way in which readers/listeners process sentences. These experiments, taken as a whole, provide evidence that while descriptions of grammatical structure are psychologically real in the sense that they are used in sentence comprehension, these structures are not used as a basis for sentence representation in memory.

In a seminal study, Sachs (1967) presented data strongly supporting the hypothesis that after the meaning of a sentence is comprehended, the exact wording and the syntactic structure are forgotten. Sachs had passages of text read to normally-hearing college students and, after various intervals of interpolated material, tested the students' recognition memory for sentences of the text. The test sentences could be different from an earlier-occurring sentence in one of the following three ways: (1) different by a semantic change; (2) different by an active/passive change; or (3) different by a formal change (e.g., "He called up Mary" vs. "He called Mary up"). For immediate testing, performance was quite accurate for all sentence change conditions. But after the presentation of as little as 27 sec. of interpolated material following the target sentence, only the semantic changes were noticed. These results have been replicated in later studies using both spoken and printed sentences and text with adults (Anderson, 1974; Begg, 1971; Begg & Wickelgren, 1974; Brewer, 1975; Fillenbaum, 1966; Johnson-Laird & Stevenson, 1970; Sachs, 1974; Wanner, 1974) and with children (Trembath, 1972).

An experiment by Bransford and Franks (1971) clearly demonstrates the tendency of subjects to recall the meaning rather than the form of sentences. Subjects in that study were presented semantically-related study sentences that contained one, two, or three propositions. In a recognition test, subjects tended to falsely recognize four-proposition sentences that they had never seen before, but which were semantically consistent with the integration of propositions across the study sentences.

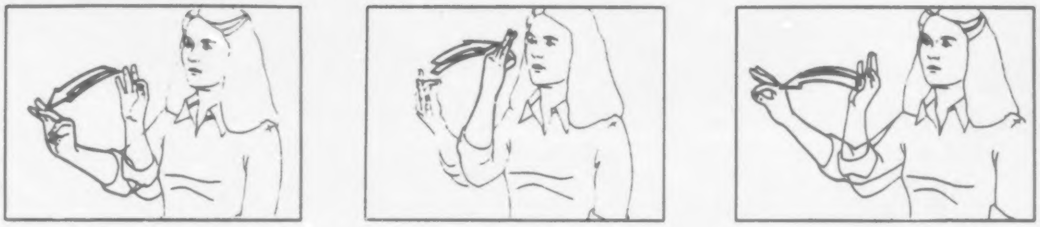
The present research is concerned with whether signers similarly abstract the meaning of ASL sentences and then retain this meaning independent of the sentence structure. At one level this experiment tests retention of sign order. The work with English readers/listeners has indicated that there is little retention of the exact word order of sentences for anything but immediate testing. It is similarly expected here that signers will not retain information about the sign order of ASL sentences in long-term memory. Work with English sentences has also shown that following sentence comprehension, readers/listeners do not retain information about the exact lexical composition of sentences. The morphological processes of ASL afford the opportunity to provide a more stringent test than is possible with English of the tendency of language users to remember the meaning of sentences independent of the lexical composition. The morphological processes of ASL, by superimposing movement patterns on basic lexical signs, strikingly alter the dynamic visual form of the sign. Of interest here is whether signers will remember the global form of a sign or whether they will decompose the complex sign into its lexical and inflectional components and remember only the meaning expressed by these components.

UNINFLECTED



PREACH (basic)

REFERENTIAL INDEXING



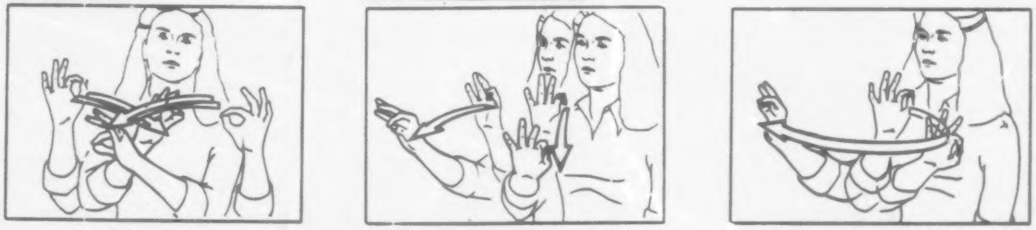
PREACH$^{\text{[Indexic: 2nd pers.]}}$
'preach to you'

PREACH$^{\text{[Indexic: 1st pers.]}}$
'preach to me'

PREACH$^{\text{[Indexic: 3rd pers.]}}$
'preach to him'

RECIPROCAL

GRAMMATICAL NUMBER



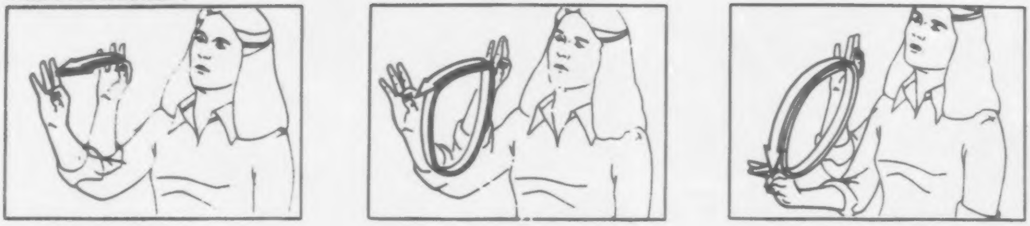
PREACH$^{\text{[Reciprocal]}}$
'preach to each other'

PREACH$^{\text{[Dual]}}$
'preach to both'

PREACH$^{\text{[Multiple]}}$
'preach to others'

Figure 1. The uninflected sign PREACH and its form under a variety of inflections.

TEMPORAL ASPECT

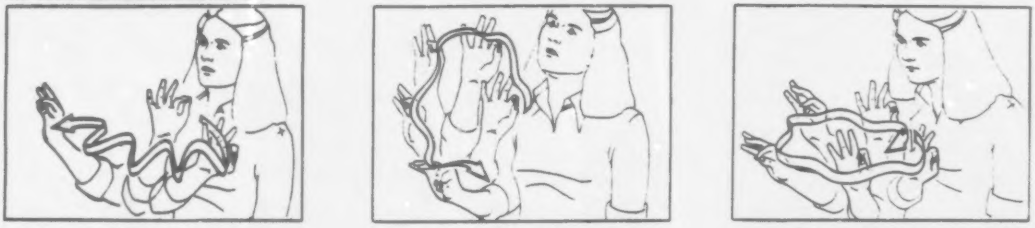


$\text{PREACH}^{[\text{Habitual}]}$
'preach regularly'

$\text{PREACH}^{[\text{Iterative}]}$
'preach over and over again'

$\text{PREACH}^{[\text{Continuative}]}$
'preach for a long time'

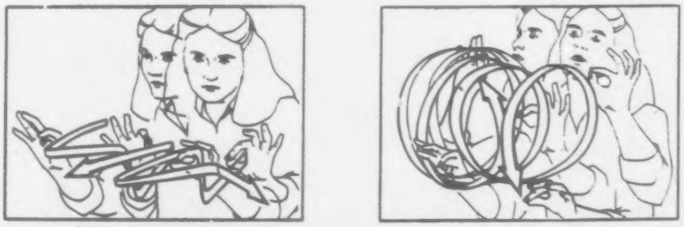
DISTRIBUTIONAL ASPECT



$\text{PREACH}^{[\text{Exhaustive}]}$
'preach to each of them'

$\text{PREACH}^{[\text{Apportionative Internal}]}$
'preach all over'

$\text{PREACH}^{[\text{Apportionative External}]}$
'preach among members of a group'



$\text{PREACH}^{[\text{Allocative Determinate}]}$
'preach to certain ones'

$\text{PREACH}^{[\text{Allocative Indeterminate}]}$
'preach to any and all'

Figure 1, continued.

The paradigm to be used is one adapted from work on memory for sentences in both written and spoken English (Begg, 1971; Begg & Wickelgren, 1974). In these paradigms, subjects are presented with several sentences. After various intervals, a test sentence is presented. Subjects are asked if that sentence is identical to an earlier-occurring one. In the present experiment, a test sentence can be changed from the target in one of four ways: Two of these changes (*Formal* and *Lexical* changes) will be paraphrases of the target sentence, preserving meaning. The other two changes (*Inflection* and *Semantic* changes) will have different meanings than the target.

METHOD

Stimulus Materials and Design

Stimuli were ASL sentences. The sentences were independent of each other in terms of grammatical structure and content. They were not related in a story context. There were three different kinds of sentences: Original, test, and filler sentences. Original sentences were the first presentation of an experimental sentence. Test sentences were the second presentation of an experimental sentence.

Fifty experimental sentence pairs were used in five testing conditions. In one condition, the original and test sentences were identical. Eighteen pairs of experimental sentences were included in this condition. The other 32 sentence pairs were equally divided among four conditions in which the test sentence was changed from the original sentence. There were eight sentence pairs in each of the following four change conditions:

a) *Formal*: Formal changes involved a change in sign order with no resultant change in meaning. An example of a Formal change is given below. Both the ASL gloss and the English translation are presented.

Original: DOCTOR NO [X:'to me'] EAT CHEESE MILK ETC. THAT.
English translation: The doctor told me not to eat dairy products.

(1) Test: CHEESE MILK ETC. THAT DOCTOR NO [X:'to me'] EAT
English translation: Dairy products are the food the doctor told me not to eat.

b) *Lexical*: Lexical changes involved pairs of sentences that had the same meaning expressed in two different ways. Thus in one sentence a particular meaning was carried by two lexical signs while in the other the same meaning was conveyed by a single inflected sign. An example of such a sentence pair is the following:

Original: SUPPOSE WINTER (MY) FRIEND [iD:'chum'] S-U-R-E OFTEN SICK PITY [X:'him'].
English translation: If it's winter my chum will surely get sick often, poor thing.

(2) Test: SUPPOSE WINTER (MY) FRIEND [iD:'chum'] S-U-R-E SICK [M:Frequentative] PITY [X:'him'].
English translation: If it's winter my chum will surely get sick often, poor thing.

The difference in form between the phrase OFTEN SICK and the modulated SICK[M:Frequentative], both of which have the same meaning, is shown in Figure 2.

c) Inflection: Inflection changes involved pairs of sentences that were different in meaning. The original sentence included inflected signs; in the test sentence, the inflections were transposed, but the order of the sign stems was maintained, resulting in a different meaning. The following sentence pair is an example in which the inflections on GIVE and INFORM are exchanged. Figure 3 illustrates this change.

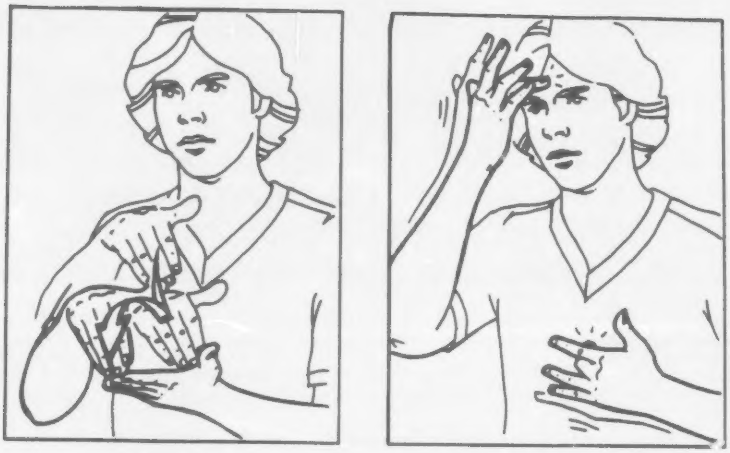
Original: MARY, JANE PAPER GIVE[N:Reciprocal], BUT NEVER INFORM[I:Multiple] THEIR WORK.
English translation: Mary and Jane give each other papers, but never inform others about their work.

(3) Test: MARY, JANE PAPER GIVE[N:Multiple] BUT NEVER INFORM[I:Reciprocal] THEIR WORK.
English translation: Mary and Jane give others papers, but never inform each other about their work.

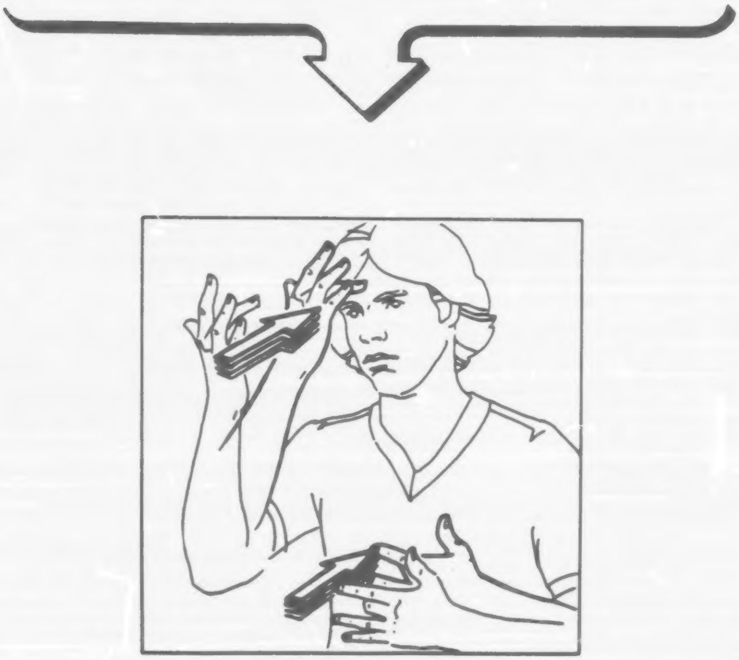
d) Semantic: For Semantic changes, the meaning of the test and original sentences was altered by a change in sign order in which two signs exchanged positions, often across clauses. An example of such a Semantic change is given here, where MATH and MUSIC are interchanged.

Original: MY WIFE TEACH [M:Continuative] MATH ALL-MORNING, AFTERNOONS TEACH [M:Habitual] MUSIC.
English translation: My wife works hard teaching math all morning, and in the afternoons she generally teaches music.

(4) Test: MY WIFE TEACH [M:Continuative] MUSIC ALL-MORNING, AFTERNOONS TEACH [M:Habitual] MATH.
English translation: My wife works hard teaching music all morning, and in the afternoons she generally teaches math.

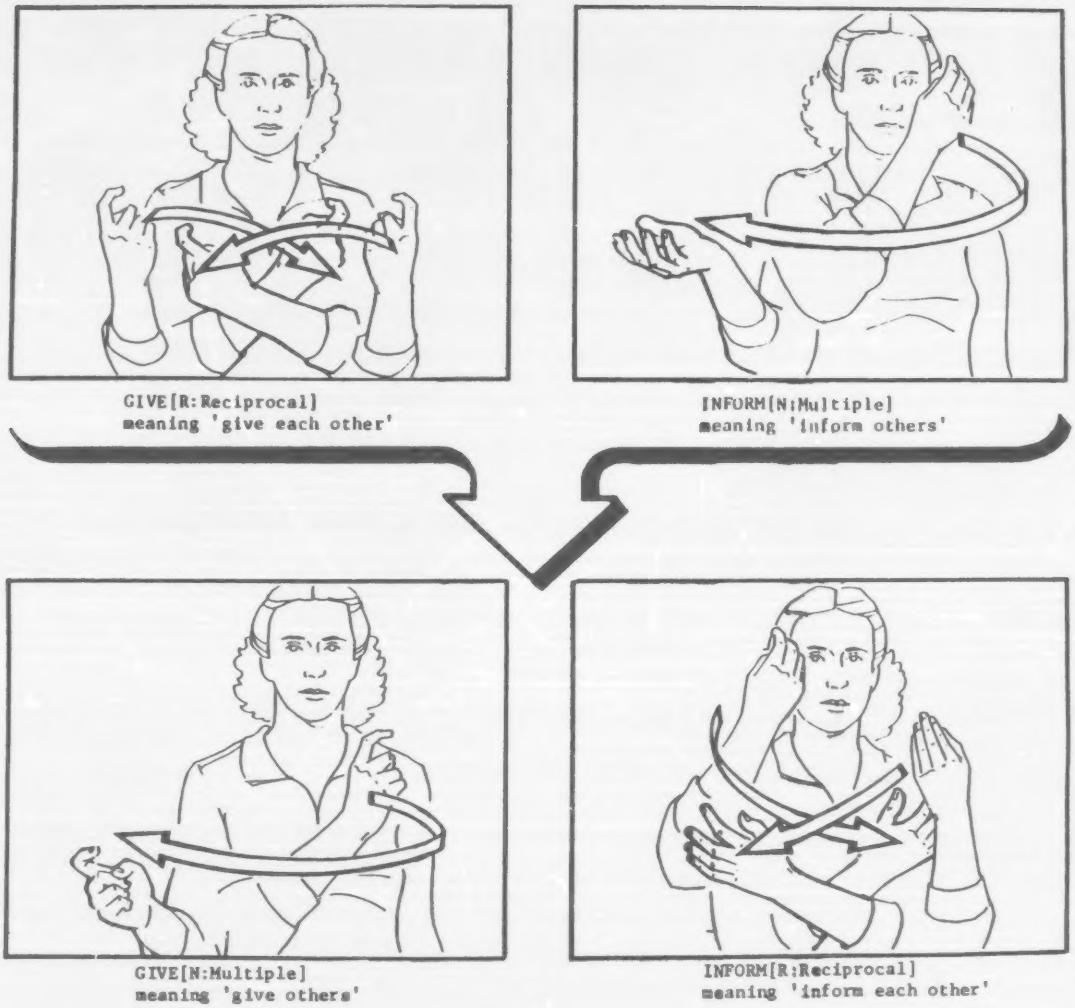


OFTEN **SICK (Basic)**



SICK$^{[M:Frequentative]}$
meaning 'often sick'

Figure 2. An example of a Lexical change. In the top panel, the meaning 'often sick' is expressed by two lexical signs: OFTEN and SICK. In the lower panel, the same meaning is expressed by the inflected sign SICK[M:Frequentative].



GIVE[R:Reciprocal]
meaning 'give each other'

INFORM[N:Multiple]
meaning 'inform others'

GIVE[N:Multiple]
meaning 'give others'

INFORM[R:Reciprocal]
meaning 'inform each other'

Figure 3. An example of an Inflection change. In the top panel, the signs GIVE and INFORM are inflected with the reciprocal and multiple inflections respectively. In the bottom panel, the same signs occur but with transposed inflections, resulting in different meaning.

There were thus five different experimental conditions as defined by the different relationships between original and test sentence pairs. In addition, there were two different time intervals tested: an immediate test and a 45 second delayed test. Sentences in the five experimental conditions were tested half the time at each interval.

Stimulus Tape

Stimulus sentences were signed by a native signer of ASL and were recorded on videotape for use during testing. Natural facial expression was included in all sentences. The hands of the signer were restored to a neutral position between sentences to indicate the end of one sentence and the beginning of another. The first three sentences presented were filler sentences.

Test sentences were indicated by a star that appeared in the upper left-hand corner of the screen at the onset of each such sentence. Following each test sentence, a blank interval lasting approximately five seconds was included to be used as a response interval.

For the immediate test, the test sentence followed the original sentence. For the delayed test, four sentences always intervened between original and test sentences. These four intervening sentences included original and test sentences as well as filler sentences. In many cases, there was also one response interval between original and test sentences. This difference in events between original and test sentences was caused by the variance in length of sentences used in the experiment. The number of intervening sentences was held constant, however, and the time of the delay interval was held constant at 45 seconds.

Instructions, signed in ASL, were recorded on the beginning of the videotape.

Procedure

Subjects were instructed that they would see several ASL sentences and that they were to pay careful attention to each. At various times test sentences would be presented and would be indicated by a star in the upper left-hand corner of the screen. For each, subjects were to decide if the test sentence was exactly the same as a sentence that had been presented previously. They were to circle YES on their answer sheet if the test sentence was the same as an earlier-occurring sentence and to circle NO if the test sentence was not exactly the same as an earlier one. Subjects were instructed that "exactly the same" meant the same signs and same sign order as well as the same meaning.

In addition, subjects were asked to make a confidence judgment about each sentence. They were to circle whether they were "VERY SURE," "SORT OF SURE," or "GUESSING" about their decision as to whether the original and test sentences were the same.

The stimulus sentences were preceded by a practice session that included four practice test sentences. All sentences included in the practice phase were simple sentences designed to illustrate clearly the nature of the procedure and to indicate that the structure as well as the meaning of the sentences would be important in the experiment. During this practice, subjects' answers were checked after each response. If a subject had answered incorrectly, the original and test sentences were shown to the subject again.

Subjects

Subjects were ten deaf volunteers recruited through the Center on Deafness at California State University, Northridge. Nine of the subjects had deaf parents and had learned ASL as a first language. The other person had grown up signing and was considered by native signers to be a skilled ASL user. There were five women and five men, mean age 24.2 years.

RESULTS

The percentage of trials on which the subjects responded that the test sentence was "different" from the original sentence is given in Table 1. For

Table 1

Mean percentage responses in which subjects responded that the test sentence was different from the original sentence. Also shown is the difference in such responses at the two time intervals. (A negative number as the difference indicates that there was more tendency to respond that the test sentence was different for delayed than for immediate testing.)

	Immediate test	Delayed test	Difference
Formal	97.5%	59.2%	38.3% *
Lexical	92.5%	67.5%	25.0% *
Inflection	90.0%	81.7%	8.3%
Semantic	90.0%	95.0%	-5.0%
Identical	28.4%	37.3%	-8.9%

(*$p < .05$)

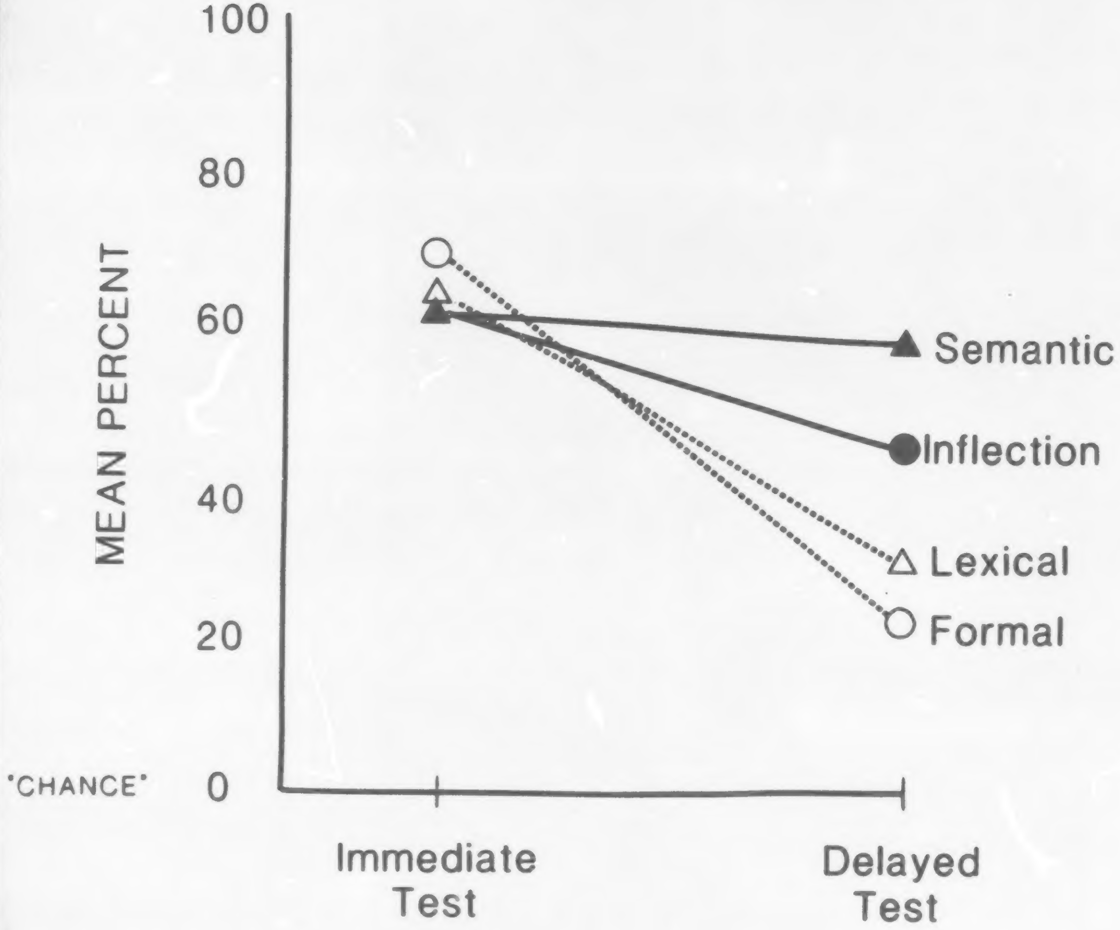


Figure 4. Percentage of "different" responses greater than "chance" for the four change conditions.

all but the Identical sentences this percentage indicates percentage correct responses. For Identical sentences there was no difference in subjects' accuracy at the two response intervals, $t(9)=1.48$, $p>.05$.

The percentage of "different" responses for the Identical sentences may be taken as an index of subjects' bias to respond that a test sentence is not the same as the original. Following Sachs, "chance" is therefore defined here as the percentage of "different" responses for Identical sentences. The results are graphed in Figure 4 as the percentage of "different" responses greater than "chance."

Analyzing the percentage of "different" responses for the four change conditions, with immediate testing it was found that subjects were equally likely to respond that the test sentence was different from the original for all four sentence types, $F(3,27)=1.00$, $MS_e=125.00$, $p>.05$. Subjects were therefore able to notice all four types of sentence changes equally well with immediate testing. They did not, however, notice the different types of sentence changes equally well with delayed testing. An analysis of variance on the four types of change conditions by intervals indicated main effects of both condition, $F(3,27)=4.22$, $MS_e=195.30$, $p<.025$, and interval, $F(1,9)=16.35$, $MS_e=339.56$, $p<.01$, that were qualified by an interaction of interval by condition, $F(3,27)=10.20$, $MS_e=176.02$, $p<.001$. Thus, while performance was generally better with the immediate test than with the delayed test, the degree to which the time interval adversely affected performance was dependent upon the condition being tested. Results of a Tukey (hsd) post hoc analysis indicated that there were significant differences between the immediate and delayed testing only for Formal and Lexical changes (those that preserved meaning) ($p<.05$). Inflection and Semantic changes, both of which changed meaning, were noticed as well with the delayed as with the immediate test ($p>.05$).

A one-way analysis of variance was performed on accuracy in the five delayed test conditions. The main effect of condition was significant, $F(4,36)=19.49$, $MS_e=248.47$, $p<.01$. Post hoc analyses were undertaken to determine the basis for this effect. Results of these tests indicated that the percentage of "different" responses was greater than "chance" for all four change conditions (Dunnett's t statistic, $p<.05$). Additional analyses indicated a distinction between meaning-preserving and meaning-changing sentence pairs: For both the Inflection and Semantic changes, subjects were more likely to respond that the test sentence was "different" than they were to respond "different" for the Formal and Lexical changes (Newman-Keuls, $p<.05$). There was no difference in the percentage of "different" responses between Semantic and Inflection changes (Newman-Keuls, $p>.05$) nor between Formal and Lexical changes (Newman-Keuls, $p>.05$). These results indicate that while subjects responded more accurately than would be expected by chance for all four types of sentence changes, their ability to notice the sentence changes was dependent on the type of change: changes of meaning (Inflectional and Semantic changes) were noticed more consistently than were meaning-preserving changes (Formal and Lexical changes).

To obtain a score for the confidence rating, subjects' responses were assigned the following numerical values: VERY SURE=3, SORT OF SURE=2, GUESSING=1. If the subject responded that the test and original sentences were identical, their confidence rating was multiplied by -1. If the subject responded that the test and original sentences were not the same, their response was given a score equal to the confidence rating. Subjects' scores on this confidence rating were then analyzed. Mean confidence ratings are shown in Table 2.

Table 2

Mean confidence ratings at both immediate and delayed testing. A positive number indicates a tendency to respond that the original and test sentences were different. A negative number indicates a tendency to respond that the original and test sentences were identical. The scoring procedure is explained in the text.

	Immediate test	Delayed test	Difference
Formal	2.88	.56	2.32 *
Lexical	2.57	.97	1.60 *
Inflection	2.30	1.77	.53
Semantic	2.40	2.67	-.27
Identical	-1.24	-.73	-.51

(*$p<.05$)

A t-test was performed on subjects' confidence in responding that the Identical test sentences were the same as an earlier-occurring sentence at the two response intervals. The nonsignificant results, $t(9)=1.69$, $p>.05$, indicated that for Identical sentences there was no difference in subjects' confidence of their responses at the two time intervals.

Scores for immediate test in the four change conditions showed no difference in confidence for the different conditions, $F(3,27)=1.71$, $MS_e=.372$, $p>.05$. Scores were then subjected to an analysis of variance on condition by interval. Results indicated main effects of both condition, $F(3,27)=5.25$, $MS_e=.533$, $p<.01$, and interval, $F(1,9)=25.02$, $MS_e=.871$, $p<.01$, that were qualified by an interaction of the two variables, $F(3,27)=11.16$, $MS_e=.587$, $p<.001$, Again, a post hoc Tukey (hsd) analysis indicated the difference in performance at the two intervals was apparent only for the Formal and the Lexical changes ($p<.05$). This pattern of results reflected the fact that with delayed testing, subjects were not confident of responding that the test sentences were different from the original sentences for the Formal and Lexical changes although they were confident of responding that the test sentences were different from the original sentences for the Inflection and Semantic changes.

In this experiment, therefore, subjects were both accurate and confident in noticing Inflection and Semantic changes even with 45 seconds intervening between original and test sentences. In contrast, sentence changes that preserved meaning (the Formal and Lexical changes) were accurately and confidently noticed only with immediate testing.

DISCUSSION

This experiment represents one of the first attempts to study ASL sentence processing. It was found that signers use syntactic structure to comprehend sentences but represent the meaning of ASL sentences in long-term memory independently of the sign order and the holistic sign structure.

As in studies with English, paraphrases were not noticed well with anything but immediate testing. This does not mean that only sentence meaning was retained. Sachs (1967) also noted this fact, stating that subjects "did have some ability to recognize the form of the sentence but that it was quite low and contrasted greatly with their memory for the semantic content of the sentence" (p. 441). The present finding that all four types of sentence changes are recognized better than "chance" is consistent with this conclusion. In particular, research has shown that when subjects know in advance that their memory for sentences is being tested (as in the case of the present study), additional information about the surface form of sentences is retained (Anderson, 1974; Begg & Wickelgren, 1974; Johnson-Laird & Stevenson, 1970). This is similar to work showing that people can remember such "trivia" as the typography of words when reading (Kolers & Ostry, 1974). Thus, it is apparently the case that people can remember many types of information about the sentences they process: However, there is strong evidence that the primary information remembered about sentences is the semantic interpretation for written and spoken sentences and, as shown here, for signed ASL sentences.

One type of paraphrase in the present study involved a change in word order. These Formal changes were not noticed with delayed testing even though, as shown in sentence pair (1), the sign order changes generally involved a topicalization change. Sentences in ASL often follow what has been referred to as a topic-comment structure. This means that the topic of the sentence, marked by a specific facial expression, occurs first, and is followed by a comment on that topic. For example, in the test sentence of sentence pair (1), the topic is "dairy products." The comment is that "the doctor told me not to eat [them]." In the original sentence for that pair, the topic of the sentence was that "the doctor told me not to (do something)" and the comment explains that the prohibited activity is eating dairy products.

Sentence topicalization in ASL is marked by an eyebrow raise and an upward head tilt (Liddell, 1977). The beginning of the comment is signaled by a relaxation of head and eyebrow position. A topicalization change was present for three of the four test sentences at each delay. The fact is that subjects did not notice even these Formal changes.

Semantic changes in this experiment, as well as Formal changes, were caused by changes in sign order. But when the sign order change also caused a meaning change, the change was noticed. Thus, it is not the sign order, per se, that was remembered, but rather it was the semantic content that was remembered.

One important aspect of the present results is that signers were shown to be representing the meaning of a signed sentence independently of its morphological composition. This was shown strikingly by the results of the Lexical change condition. Subjects remembered the meaning of the sentence but after a brief delay did not remember the form in which the meaning was conveyed. Thus, for example, subjects remembered the meaning 'frequently sick' but did not remember if the actual sentence contained a two-sign phrase OFTEN SICK or a single inflected sign, SICK[M:Frequentative].

This finding agrees nicely with work by Poizner, Newkirk, Bellugi, and Klima (1981) on short-term memory for lists of inflected signs. In that study, subjects saw short lists of inflected signs and were asked to recall the list by signing the items immediately after each list. Recall errors revealed that subjects recalled the base sign and its inflection independently. Thus, signers decomposed the sign into its meaning components and did not retain the exact form of the sign.

Notice, however, a striking difference in this work on sentence memory and the results from the short-term memory paradigm of Poizner et al. (1981). For the lists of inflected signs in their study, subjects confused which inflections were superimposed on which basic signs. But this was not true when the inflected signs were put in the meaningful sentence context of the present experiment. The high recognition accuracy for Inflection changes, even at longer intervals, indicates that people did not make this confusion. Results from the Lexical change condition further show that this lack of confusion was not a result of signers remembering the exact form of the sign presented. Rather, this accurate performance in the Inflection change condition can be attributed to the fact that switching the inflections changed the meaning of the test sentence. Thus, signers were able to reject the test sentences for Inflection changes, not because they noticed that the sign forms were changed but because they noticed that the meanings of the test sentences were changed.

In recent years a picture has been emerging in which basic cognitive processes constrain the structure of signed and spoken languages so that underlying structure for languages in both modalities exhibit many similarities (Bellugi, 1980). The present work on sentence memory suggests that in signed language as in spoken languages, the meaning of a sentence is abstracted and the structural mechanism by which this meaning is conveyed is not retained in the long-term memory representation of the sentence.

REFERENCES

Anderson, J. R. Verbatim and propositional representation of sentences in immediate and long-term memory. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 149-162.

Battison, R. Phonological deletion in American Sign Language. Sign Language Studies, 1974, 5, 1-19.

Begg I. Recognition memory for sentence meaning and wording. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 176-181.

Begg, I., & Wickelgren, W. A. Retention functions for syntactic and lexical vs. semantic information in sentence recognition memory. Memory & Cognition, 1974, 2, 353-359.

Bellugi, U. Clues from the structural similarity of sign and spoken language. In U. Bellugi & M. Studdert-Kennedy (Eds.), Signed and spoken language: Biological constraints on linguistic form. Weinheim: Verlag Chemie, 1980.

Bellugi, U., Klima, E. S., & Siple, P. Remembering in signs. Cognition, 1975, 3, 93-125.

Bransford, J. D., & Franks, J. J. The abstraction of linguistic ideas. Cognitive Psychology, 1971, 2, 331-350.

Brewer, W. F. Memory for ideas: Synonym substitution. Memory & Cognition, 1975, 3, 458-464.

Fillenbaum, S. Memory for gist: Some relevant variables. Language and Speech, 1966, 9, 217-227.

Frumkin, B., & Anisfeld, M. Semantic and surface codes in the memory of deaf children. Cognitive Psychology, 1977, 9, 475-493.

Johnson-Laird, P. N., & Stevenson, R. Memory for syntax. Nature, 1970, 227, 412.

Klima, E., & Bellugi, U. The signs of language. Cambridge, Mass.: Harvard University Press, 1979.

Kolers, P. A., & Ostry, D. J. Time course of loss of information regarding pattern analyzing operations. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 599-612.

Liddell, S. K. An investigation into the syntactic structure of American Sign Language. Unpublished doctoral dissertation, University of California, San Diego, 1977.

Newkirk, D., Klima, E. S., Pedersen, C. C., & Bellugi, U. Linguistic evidence from slips of the hand. In V. Fromkin (Ed.), Errors in linguistic performance: Slips of the tongue, ear, pen, and hand. New York: Academic Press, 1980.

Poizner, H., Newkirk, D., Bellugi, U., & Klima, E. S. Representation of inflected signs from American Sign Language in short-term memory. Memory & Cognition, 1981, 9, 121-131.

Sachs, J. S. Recognition memory for syntactic and semantic aspects of connected discourse. Perception & Psychophysics, 1967, 2, 437-442.

Sachs, J. S. Memory in reading and listening to discourse. Memory & Cognition, 1974, 2, 95-100.

Stokoe, W. C., Jr., Casterline, D., & Croneberg, C. A dictionary of American Sign Language. Washington, D.C.: Gallaudet College Press, 1965.

Trembath, A. A. Comparison of sensitivity to the surface and deep structure of sentences in children. Language and Speech, 1972, 15, 51-57.

Wanner, E. W. On remembering, forgetting, and understanding sentences. The Hague: Mouton, 1974.

FOOTNOTE

[1] Words in capital letters represent English glosses for ASL signs. The gloss represents the meaning of the basic uninflected form of a sign. A bracketed symbol following a sign gloss indicates the grammatical process the sign has undergone (X: for referential indexing; M: for modulation for temporal aspect or focus; N: for numerosity inflection; D: for derivational process; iD: for idiomatic derivative). The symbol may be followed by a specification of the inflectional process or by the meaning of the inflected form. For example, GIVE[N:Exhaustive] and GIVE[N:'to each'] are alternative ways of representing the same inflectional process. The solid bar above specific parts of a sentence indicates sentence topicalization.

PERCEPTION OF NASAL CONSONANTS WITH SPECIAL REFERENCE TO CATALAN

Daniel Recasens+

1. INTRODUCTION

In this paper I study the role that different place cues play in the recognition of nasal stops. I claim that their perceptual relevance is strongly dependent on how they are related at the articulatory and acoustic levels and, essentially, on the nature of the process of speech perception itself. I show that this is the case by investigating experimentally interactive perceptual effects between transitions and murmurs in the recognition of final unreleased alveolar [n], palatal [ɲ] and velar [ŋ] after [a] in Catalan, using synthetic speech stimuli.[1] Special emphasis is given to the cues for the palatal nasal.

I proceed first to investigate what acoustic properties of the signal can be shown to convey place information by looking at a large amount of production and perceptual data on nasal murmurs and formant transitions. The role of releases in the process of place identification for nasals is also taken into account. A consideration of other cues besides formant transitions seems highly advisable. In an early perceptual experiment with synthetic speech (Liberman, Delattre, Cooper, & Gerstman, 1954) it was found that, in contrast with initial non-nasal stops, final nasal consonants ([m], [n], [ŋ]) after different vowels were properly identified only 55% of the time for stimuli with appropriate transition endpoints and a cross-category fixed nasal murmur. Results from more recent experiments, both with synthetic (Garcia, 1966, 1967a, 1967b; Hecker, 1962; House, 1957; Nakata, 1959) and with natural (Henderson, Note 1; Malécot, 1956; Nord, 1976) speech stimuli have shown that not only formant transitions but also murmurs and releases are cues to place of articulation for nasal consonants. As will be shown, experimental data from the literature on speech perception suggest that all these cues for nasals ought to be considered as interdependent and, therefore, need to be taken into account in models of the perceptual evaluation of place cues.

+Also University of Connecticut, Storrs.

Acknowledgment. I acknowledge very gratefully the assistance of Ignatius G. Mattingly at all stages of this investigation and, especially, in the preparation of the manuscript. I would also like to thank Arthur S. Abramson, Louis Goldstein, Alvin Liberman, Bruno Repp, and the Faculty of the Department of Linguistics of the University of Connecticut for their relevant comments. The research was supported by NICHD Grant HD01994 and BRS Grant RR05596 to Haskins Laboratories. An abstract of this paper was presented at the 101st Meeting of the Acoustical Society of America (Journal of the Acoustical Society of America, 1981, 69, S83). I am also appreciative of Margo Carter's help with the figures.

One must also consider the place cues for [ɲ] within the overall nasal set. It will be pointed out in what way their being taken into consideration at the analysis and synthesis levels affects theoretical considerations about perceptual relevance of transition patterns advanced by other scholars in earlier experiments that did not account for [ɲ].

The allophonic system of Catalan nasals in absolute final position is adequate to test these hypotheses with synthetic speech experiments. It consists of unreleased [m], [n], [ɲ], [ŋ] and allows analysis of the perceptual effects of transitions vs. murmurs with special reference to place cues for [ɲ]. In my experimental paradigm, which differs from approaches taken previously by other investigators, complete patterns of synthetic transitions and murmurs are directly based on real speech utterances and combined reciprocally in perceptual continua for all the different place categories. Perceptual results are related to production data on nasals collected from Catalan speakers and speakers of other languages, and discussed in the light of the literature. Evidence for a complementary perceptual influence of transitions and murmurs consistent with parallel effects observable for both cues at the acoustic level is reported. This and other findings argue for some form of motor theory (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967) that refers to the unitary articulatory gesture to account for the perceptual processing of dynamic acoustic cues in syllables ending with nasal stops. An integration model similar to that of Dorman, Studdert-Kennedy, and Raphael (1977) for non-nasal stops is proposed to account for the perception of nasals after [a] as well as other vocalic nuclei.

Data for nasal consonants in syllable-final position are taken into consideration because of the fact that the perceptual effectiveness of murmurs in place recognition in this position is known to be considerably higher than in syllable-initial position (Malécot, 1956; Nord, 1976). Open vowels [a], [æ] have been chosen for analysis and discussion since the perception of consonantal nasalization improves with [a] vs. [i], [u] (Ali, Gallagher, Goldstein, & Daniloff, 1971; Mártony, 1964; Zee, 1981). Also [ŋ], which happens to be harder to identify, in general, than other nasal consonants (Garcia, 1966, 1967a,b; Malécot, 1956; Ohala, 1975), can be recognized with [a] quite easily when synthesized (Hecker, 1962) or presented in natural speech (Wang & Fillmore, 1961), but rather poorly when it happens to be contiguous to [i], [u]. In section 4, I will also refer to the interactive effects of nasal cues and other vowel nuclei.

2. CUES FOR NASAL CONSONANTS: PERCEPTUAL RELEVANCE, ARTICULATORY AND ACOUSTIC CHARACTERISTICS

2.1. Manner Cues

Certain well-defined spectral characteristics of nasal murmurs mark nasal consonants as a class, independent of place of articulation and the adjacent nasalized vowels (Delattre, 1958, 1968; Fant, 1960; Fujimura, 1962; Fujimura & Lindqvist, 1970; Hattori, Yamamoto, & Fujimura, 1958; Mattingly, 1968). Formant transitions, on the other hand, are essentially place markers; in fact, as shown below, only the first formant transition contributes effectively to manner identification. In the following paragraphs I refer to those spectral characteristics, their general articulatory correlates and perceptual relevance, in preparation for discussing not only differences in murmur patterns for nasals of different place of articulation but also those experimental paradigms concerned with place identification that make use of fixed or slightly variable murmurs:

a. First formant (N1), at around 250-300 Hz, with higher intensity than the upper spectral regions, dependent on a large internal cavity size (pharynx and nasal subsystems) behind the tongue constriction because of nasal coupling. According to Delattre (1958), the intensity level of N1 and the other spectral regions of the murmur is around 6 dB (N1) and 15 dB (N2, N3, N4...) lower than for a normal non-nasalized vowel. It seems to be the most important class cue for nasal consonants, in contrast with the negligible perceptual role of the frequencies of higher nasal formants (Delattre, 1968).

b. Presence of an antiformant (NZ), varying in frequency with place, according to the size of the mouth cavity behind the tongue constriction, which acts as a shunt. It seems to convey mainly place information.

c. Concentration of formants (N2, N3, N4...) between 300-4000 Hz, with large bandwidth (BW) values, mainly due to the large surface area of the nasal cavities and the dissipative energy losses originated within them. The small perceptual significance of those formants seems to result not only from their low intensity level with respect to N1 (especially N2, often absent, as reported by Fant, 1962, and Weinstein, McCandless, Mondshein, & Zue, 1975) but also from their spectral variability. In synthetic speech experiments, the following ranges of fixed frequency values have been found effective in realizing the nasal murmur when place is held constant: N2: around 1000-1150 Hz; N3: around 2000-2500 Hz (Delattre, 1954; Liberman et al., 1954; Massone, 1980, for Argentinian Spanish; Miller & Eimas, 1977).

The value for N2 has been proved to be dependent on the size of the narrow velar passage to the nasal cavity (Bjuggren & Fant, 1964). The significance of an N3 around a "typical" 2200 Hz area has also been pointed out by Fant (1962); this resonance seems to be chiefly dependent on the characteristics of the pharynx cavity (Fant, 1960; Fujimura, 1962). In line with these observations, De Mori, Gubrynowicz, and Laface (1979) have recently proposed the automatic interpretation of any frequency concentration between 2000-2800 Hz as N3 and of the first nasal formant below it as N2 as a speech recognition rule for identification of nasal consonants.

There are also available data on optimal formant bandwidth values for consonantal nasality. Thus Mártony (1964) has stressed the perceptual relevance of an N2 bandwidth value around 250 Hz, given an optimal N1 value at around 100-150 Hz. Such an N2 bandwidth is close to frequencies chosen as the most favorable for nasal perception by Nakata (1959) (N2: 200 Hz; N1: 300 Hz) and Pickett (1965) (N1, N2, N3: 180 Hz; N4: 300 Hz).

d. Overall lower intensity level than vowels. House (1957) assigned to the murmur an intensity 8 dB lower than that appropriate for [i].

Other manner cues besides nasal murmurs need to be accounted for:

a. Vowel nasalization, taken into consideration in experiments with synthetic speech for a 100 msec (Haskins Laboratories QPR 13, 1954, Appendix 2) and 20 msec (Miller & Eimas, 1977) overlapping period between vowel and consonant. It can be simulated by replacing F1 by two formants (N1, F1) and an antiformant (NZ) and by increasing gradually and monotonically N1 and NZ values from 0 (absence of oronasal coupling) to 600 Hz for N1 and 650 Hz for NZ (Fujimura & Lindqvist, 1970, Figure 13) or 660 Hz for N1 and 700 Hz for NZ (Fant, 1960) (presence of a small degree of oronasal coupling). It would be desirable to reproduce the effect of higher extra pole-zero concentrations of the nasalized vowel and to reach a better understanding of the configuration of frequency continuities and discontinuities between murmur and vowel formants at the closure onset in order to find out to what extent they help to identify consonantal nasality. A continuity between vowel formants and nasal poles (F1-N1, F2-N3, F3-N4...) (Fant, 1960) is confirmed by my considerations on analysis and synthesis of nasals (section 3.2.). Moreover, Takeuchi, Kasuya, and Kido (1975) have shown that an uninterrupted pole excursion between vowel and nasal consonant, with the addition of some nasality parameter that represents the amount of spectral difference between nasal and vowel spectra, can be regarded as a valid cue for detection of nasals as a class.

b. Nasalized releases following the nasal murmur, different from non-nasal stop releases in presenting low-frequency masking (Blumstein & Stevens, 1979).

c. F1 transitions, generally negative but less so than for non-nasal stops. This differential acoustic cue has perceptual relevance for nasals as a class (Fant, 1967; Mattingly, 1968; Miller & Eimas, 1977).

2.2. Place Cues

A. Nasal murmur

It has been suggested that nasal murmurs also play a relevant role in identification of different places of articulation. Thus, experimenters at Haskins Laboratories emphasized very early the polyvalent nature of formant transitions and nasal murmur characteristics in the process of place discrimination among nasal consonants (Cooper, Delattre, Liberman, Borst, & Gerstman, 1952). As Fant (1980) has proposed recently: "The base rule stating that stationary segments signal the manner and transitions the place of articulation has more exceptions than one might expect. Thus the nasal murmur of [m], [n], and [l] may contain strong place cues..." (p. 14). In the light of such observations, I will argue for consistent place cues in the nasal murmur portion by comparing relevant production and perceptual data.

Table 1 presents mean and extreme frequency values for N1, N2, N3, N4 and NZ in production data from male speakers of different languages (Czech, German, English, Hungarian, Polish, Russian, and Swedish) reported by several authors. A summary of the results, included at the bottom part of Table 1, shows that formant patterns for [ɲ] and [ŋ] are very similar except in the case of N4, which is higher for [ŋ] than for [ɲ]; [m] and [n] present lower

Table 1

Pole-zero analysis frequency values (in Hz) corresponding to the nasal murmurs of [m], [n], [ɲ], [ŋ] in Czech (9), German (9), English (4), Hungarian (8), Polish (1, 5, 6, 7), Russian (2, 9) and Swedish (3) (male speakers). A summary of values is also included.

	N1	N2	N3	N4	NZ	References
[m]	300-340	850-900	1150-1500	1700-2000	—	(1)Dukiewicz (1967)
	250	800	1000	2000	550	(2)Fant (1960)
	100-350	875-1300	2000-2500	—	—	(3)Fant (1973)
	280-290	950-980	1360	1950	750-1250	(4)Fujimura (1962)
	100-400	800-1200	2150-2500	2800-3000	—	(5)Jassem (1964)
	300	750	1300	2000	800	(6)Jassem (1973)
	300	850	1300	1850	1100	(7)Kacprowski (1963)
	—	850	2000-2600	—	—	(8)Magdics (1969)
	—	600-800	1300-1900	1800-2900	—	(9)Romportl (1973)
[n]	310-430	900-1000	1500-1800	1900-2350	—	(1)
	250	800	1200	2000	1800	(2)
	100-300	1000-1750	2400-2900	—	—	(3)
	300	1000	1450	2000	1450-2700	(4)
	100-400	1700-2000	2150-2500	2800-3000	—	(5)
	330	1200	1750	2050	1450	(6)
	320	900	1500	2050	1750	(7)
	—	1000-1600	2050	—	—	(8)
	—	830	1900	2200	—	(9)
[ɲ]	310-430	1000-1200	1500-1600	2000-2300	—	(1)
	250	800	1900	2600	2200	(2)
	100-400	1100	2200-2500	2800-3000	—	(5)
	345	1225	2100	2750	2750	(6)
	340	1000	1650	2180	2400	(7)
	—	1500-2000	—	—	—	(8)
	—	900-1100	2200-2500	2850	—	(9)
[ŋ]	320-440	900-1300	1500-2200	2000-2550	—	(1)
	300	1000	2200	2900	—	(2)
	100-500	1000-1250	1625-2400	3200-4000	—	(3)
	250-400	950-1150	1700-2200	2300-3000	above 3000	(4)
	100-400	900-1100	2200-2500	2800-3000	—	(5)
	350	1100	1900	2750	above 3800	(6)
	350	1050	1800	2450	—	(7)
	—	750-1500	2300-2600	—	—	(8)
	—	800	2000	2600	—	(9)

Summary

[m]	100-400	600-1300	1000-2600	1700-3000	550-1250	(Extremes)
	285	795	1610	2090	860	(Mean)
[n]	100-430	600-2000	1200-2900	1900-3000	1450-2700	(Extremes)
	300	1115	1790	2190	1770	(Mean)
[ɲ]	100-430	800-2000	1500-2500	2000-3000	2200-2400	(Extremes)
	320	1140	1920	2450	2450	(Mean)
[ŋ]	100-500	750-1500	1500-2600	2000-4000	above 3000-3800	(Extremes)
	330	1060	2015	2650	above 3400	(Mean)

formant values, those for [m] being even lower than those for [n]. Data on N1 frequency values present a succession [ŋ]>[ɲ]>[n]>[m] presumably related to cross-category differences in size of the coupling section at the velar pharyngeal passage and pharyngonasal tract size; although the differences in those frequency values are small, it should be pointed out that they can be found in data reported by some researchers in Table 1.[2] Data on NZ frequency values give an arrangement [ŋ]>[ɲ]>[n]>[m] that is consistent with NZ dependence on the oral cavity size behind the tongue constriction. (Also for Japanese [m] vs. [n], see Hattori et al., 1958). The summary also shows frequency proximity of NZ and some specific formant, depending on place category: [m] (N2), [n] (N3), [ɲ] (N4), [ŋ] (higher than N4) (Fujimura, 1963). A way to look at the distinctive rôle of NZ is by considering its position with respect to two adjacent formants (a pole-zero cluster) at different frequency regions according to nasal categories of different place of articulation (Fujimura, 1963). According to data in Table 1, when values reported by different investigators are accounted for, such a classification procedure can be said to hold strictly for [n] (N3 between N4) and [ŋ] (above the general low formant structure, even above N6, according to Kacprowski & Mikiel, 1965), but not for [m], whose antiresonance can lie between N1 and N2 or N2 and N3, nor [ɲ], whose antiresonance can lie between N3 and N4 or N4 and N5. This variability seems to be partly related to differences in vowel context.

Information about perceptual relevance of released vs. unreleased murmurs as place cues is provided by experiments in which they were presented for identification alone or directly attached to the vocalic portion with no formant transitions or release. I refer to results obtained for open vowels [a], [æ] with labial, alveolar, palatal, and velar nasal murmurs in final position.[3]

In experiments with natural English speech, released murmur segments for [m] were categorized quite accurately (80%-100%) whether presented in isolation or following [æ] without formant transitions (Malécot, 1956). Henderson (Note 1) found for the overall vocalic set a higher accuracy in place identification for [m] without transition or release (about 75%) than for [n], [ŋ] (65%-75%); following [a], in the absence of formant transitions, the [m] murmur identification was always higher than 90%, independently of the presence or absence of release. Identification scores for synthetic [m] murmurs in isolation with American English subjects are reasonably consistent with results with natural speech (65%-85%) (House, 1957; Nakata, 1959). Moreover, Manrique de, Gurlekian, and Massone (1980) report in experiments with natural Argentinian Spanish speech that not only isolated [m] murmurs but also isolated [n] murmurs give a higher percentage of [m] than [n] identifications.

The syllable [an], with no transitions or release, is identified 50%-60% of the time (Henderson, Note 1); with release, the average rises to 97% in Henderson's experiment but not in Malécot's (50%). Both naturally-spoken (Malécot, 1956) and synthetic (House, 1957; Nakata, 1959) [n] murmurs presented in isolation give the same 50%-60% effect. [n] murmurs not heard as [n] tend to be heard as [m], as in Manrique de et al. (1980).

Dukiewicz (1967) has shown that [ɲ] murmurs presented in isolation to Polish speakers elicit no [ɲ] judgments. Correct recognition of [ɲ] improves 30% when murmur is presented with its corresponding onset.

Isolated [ŋ] murmurs are less well identified than those for [n] and [m] (Malécot, 1956: 12%; House, 1957: 62%; Nakata, 1959: 41%); they tend to be interpreted mainly as [m] but also as [n]. According to Malécot, released [æŋ] murmur without transitions is very poorly identified (34%) and confused mainly with [æn]. Henderson also finds [ŋ] murmur to be a poor indicator with vowel [ɛ] (45%); however, for [aŋ] without transitions, responses rise to 74% (unreleased murmur) and 92% (released murmur).

Frequency values chosen in synthetic speech experiments (English, Italian, Polish) with fixed and variable murmur patterns are another useful indirect source of reference in the investigation of murmur structures for nasals of different place of articulation (Table 2). While experiments with fixed patterns account exclusively for class properties of consonantal nasality (section 2.1.c.), these variable patterns capture N1 and NZ dynamics correctly to simulate different place properties but reproduce only to some degree the complex formant structure observed for different nasal consonants (see Table 1). A good approximation to NZ can be obtained by replacing it, together with the two surrounding poles, by a single formant having an appropriate bandwidth. This method may explain good perceptual results obtained by Nakata (1959) and Kacprowski and Mikiel (1965) with initial synthesis values listed in Table 2.

A comparison of the production and perception data given above allows us to infer presumable place cues for nasal consonants. Thus, the remarkable importance of released or unreleased [m] murmur in place identification can be related to the particular low N1 and NZ frequency values within an overall low murmur spectrum. Such spectral configuration has been reported to differ consistently and strikingly from that of the other members of the nasal set (Table 1; Malécot, 1956; Romportl, 1973). The perceptual importance of NZ for [m] vs. [n] has been reported by De Mori et al. (1979) and that of a strong low spectrum component about 1000-1500 Hz in the case of [m] vs. a high concentration about 2300 Hz for [n], [ŋ] by Delattre (1958).

Intraspeaker and interspeaker inconsistency for murmurs other than [m] has been noted in different languages (Malécot, 1956; Delattre, 1958; Romportl, 1973). As already stated, for [a] and [æ], while [ɲ] murmur seems to carry very little place information, [ŋ] murmur and [n] murmur provide important place information in the unreleased case and are identified quite consistently when released. The perceptual distinction between [ŋ] vs. [n] could be well cued by high vs. low N1 and absence vs. presence of NZ at the central region of the spectrum. Accordingly, a higher N1 value for [ŋ] (240 Hz) than for [m], [n] (180 Hz) was reported to help perceptual place identification (Haskins Laboratories QPR 11, 1954). NZ's being above 3000 Hz or absent in the case of [ŋ] is consistent with Ohala's (1975) observation that its perceptual effectiveness is presumably severely attenuated because of that high frequency location. High perceptual distinctiveness and, at the same time, similarity of murmur spectra (high N1 and NZ, similar N2-N3-N4 configuration) between [ŋ] and [ɲ] accord well with the strong role of [ɲ] transitions in [ɲ] identification (section 2.2.B.).

With respect to acoustic aspects other than spectral frequency, it has been found that N1, N3 bandwidths for [ŋ] are greater than those for [m], [n] and that [n] presents a very high degree of N2 damping with respect to [m]

Table 2

Pole-zero frequency values (in Hz) corresponding to the nasal murmurs of [m], [n], [ɲ], [ŋ], assigned to synthetic speech stimuli in experiments with English (1, 2, 3), Italian (5) and Polish (4). A summary of values is also included.

	N1	N2	N3	N4	NZ	References
[m]	250	1360	2200	—	1550	(1)Hecker (1962)
	180	1000	2250	—	750	(2)House (1957)
	200	1100	2500	—	—	(3)Nakata (1959)
	200	1100	2500	—	—	(4)Kacprowski & Mikiel (1965)
	300	1100	2000	—	—	(5)Ferrero et al. (1977)
[n]	250	1300	2000	—	4000	(1)
	240	1100	2300	—	3500	(2)
	200	1700	2500	—	—	(3)
	200	1700	2800	—	—	(4)
	300	1800	2800	—	—	(5)
[ɲ]	200	2300	2800	—	—	(4)
	400	2300	3000	—	—	(5)
[ŋ]	330	1300	2200	—	—	(1)
	200	2300	2500	—	—	(2)
	260	900	2300	—	above 5000	(3)

Summary (Extremes)

[m]	180-300	1000-1360	2000-2500	—	750-1550
[n]	200-300	1100-1800	2000-2800	—	3500-4000
[ɲ]	200-400	2300	2800-3000	—	—
[ŋ]	200-330	900-2300	2200-2500	—	above 5000

(Fujimura, 1962). High N1 damping for [ŋ] has also been reported by other investigators (Fant, 1973; Kacpranski, 1965). No evidence about the perceptual significance of these differences is available. The perceptual influence of coarticulation effects due to the adjacent vowels upon place identification of a nasal consonant has been shown in experiments on automatic speech recognition (De Mori et al., 1979); such influence is to be expected from significant variations of the pole-zero structure of the nasal murmur according to the vocalic environment. Moreover, the perceptual importance of relative amplitude levels at different spectral areas should be investigated. In this respect it is highly plausible that, according to data reported by Eek (1972), energy minima at typical NZ frequency zones (Section 2.2.A.) convey relevant perceptual information in distinguishing nasal categories of different place of articulation.

B. Formant Transitions

In Table 3 I present analysis frequency values for F2-F3 transition endpoints as well as positive (+), steady (=) and negative (-) transition ranges according to data on syllables with [a] and [m], [n], [ɲ], [ŋ] reported by several authors.[4] The four languages that have [ɲ] have been chosen are Hungarian, Italian, Polish, and Russian.

According to the summary presented at the bottom of Table 3, F2 for [ɲ] does not overlap in frequency with other categories and shows a constant (positive) transition direction. In fact, a 250 Hz separation minimum between F2 values for [n] and [ɲ] is a good indication of high F2 distinctiveness for [ɲ]. Such a difference was also found for Czech and Russian by Romportl (1973) ([n]: 1100-1300 Hz; [ɲ]: 1600-1800 Hz).

On the other hand, F2 transition values for [ŋ] overlap significantly with those for other nasals, namely [m] and [n], being in the central part of the overall range of endpoint frequency values for the consonantal set. Variability of F2 transition values with [a] from slightly rising to steady and slightly falling has also been noted for the velar non-nasal correlates [g], [k] (Fischer-Jørgensen, 1954; Halle, Hughes, & Radley, 1957; Potter, Kopp, & Green, 1947). See, however, Dalby & Port, 1980, who found quite strongly positive F2 ranges). Flat transitions were found for English [aŋ] by Green (1959).

Predictions of acoustic theory of speech production support F2-F3 values reported in Table 3. Thus, a comparison with one of Fant's nomograms (1960, p. 84) shows that, while given frequencies for labial and alveolar nasals match well with constriction points located near the lip opening area, formant frequencies for [ŋ] correspond to a constriction place at about 9 cm from lip opening and those for [ɲ] to a constriction at about 4 cm. Negative transitions for [m] are due to a complete labial constriction. On the other hand, a more forward constriction point for [n], [ɲ] than for [ŋ] causes F2 and F3 values for [a] to increase as a result of a decrease in front cavity size. Higher frequency for palatals than alveolars is presumably related to a greater increase in the conductivity index of the internal resonator neck: palatographic evidence for [ɲ] in different languages often shows alveolar contact, as for [n], plus several degrees of prepalatal and/or dorsopalatal contact.

Table 3

Analysis frequency values (in Hz) corresponding to F2-F3 transitions in syllables with [a] and [m], [n], [ɲ], [ŋ] in Hungarian (3), Italian (5, 6), Polish (1, 2) and Russian (4). A summary of values is also included.

	Endpoints		**Ranges**		**Direction**		**References**
	F2	**F3**	**F2**	**F3**	**F2**	**F3**	
[m]	1000-1100	2850-3050	-50/-250	+50/+100	-	+	(1)Dukiewicz (1967)
					-/=	—	(2)Jassem (1962,1964)
	$\overline{1050}$	$\overline{1695}$	$\overline{-380}$	$\overline{-500}$	-	-	(3)Magdics (1969)
	800	2150	-500	-50	-	-	(4)Fant (1960)
	1200	2340	-200	-140	-	-	(5)Vagges et al. (1978)
[n]	1250-1500	2800-2900	+50/+200	-150/+200	+	-/+	(1)
					+	—	(2)
	$\overline{1650}$	$\overline{1650}$	$\overline{+220}$	$\overline{-540}$	+	-	(3)
	1400-1500	2250	0	-50	=	-	(4)
	1425	2665	+55	+40	+	+	(5)
	1450-1500	2650	+50	-50	=/+	-	(6)Ferrero et al. (1979)
[ɲ]	1900-2300	2500-3150	+500/+1000	-450/+550	+	+/-/=	(1)
					+	+	(2)
	$\overline{1970}$	$\overline{2032}$	$\overline{+495}$	$\overline{-195}$	+	-	(3)
	2050	2850	+550/+650	+550	+	+	(4)
	2140	2990	+825	+475	+	+	(5)
	2200	3200	+650	+350	+	+	(6)
[ŋ]	900-1400	2900-3200	-50/-200	-100/+600	-	-/+	(1)
	—	—	—	—	-/+	-	(2)
	—	—	—	—	=/-	—	(3)

Summary

	Endpoints		**Ranges**		**Direction**	
	F2	**F3**	**F2**	**F3**	**F2**	**F3**
[m]	800-1200	1695-3050	-50/-500	-500/+100	-(=)	-(+)
[n]	1250-1650	1650-2900	0/+220	-540/+200	+(=)	-(+)
[ɲ]	1900-2300	2030-3200	+495/+1000	-450/+550	+	+(-, =)
[ŋ]	900-1400	2900-3200	-50/-200	-100/+600	-(=, +)	-(+)

Besides conveying manner information, F1 transition values seem to contribute to place identification. Thus, while the [ɲ] transition is extremely negative (Fant, 1960; Vagges, Ferrero, Caldognetto-Magno, & Lavagnoli, 1978), that for [ŋ] is usually only slightly negative and can even be positive (Dukiewicz, 1967). For [a] followed by [ɲ], the negativity is due to an important increase of back cavity size and a noticeable increase of vocal tract constriction (Delattre, 1951; Fant, 1960). The slightly negative F1 excursion between [a] and [ŋ] is related to a smaller increase in pharynx cavity size.

Available data on perceptual experiments with natural speech and synthetic speech give information about the relevance of formant transitions in place recognition. According to Henderson (Note 1), sequences of [a] followed by a nasal consonant, without nasal murmur and with or without final release, give 60% of [m] responses when presented with [m] transitions and 80% of [n] responses with [n] transitions but only 15%-30% of [ŋ] responses with [ŋ] transitions in favor of a majority of [an] judgments. For [aɲ], with natural Polish speech, Dukiewicz (1967) found that transitions compensated for the negligible place information conveyed by murmurs (section 2.2.A.).

In Table 4 I present transition frequency values reported in perceptual tests with English-speaking subjects for synthetic [m], [n], [ŋ] and vowels [a], [æ]. Unlike the murmurs, there is much available perceptual data on transition cues. A comparison of values for [a] in Tables 3 and 4 shows that F2 frequency values categorized as [ŋ] in synthetic speech experiments (1920-2300 Hz) correspond exactly to analysis values for F2 of [ɲ] while analysis values for F2 of [ŋ] correspond to values categorized as [m] and [n] in experiments with synthetic speech. It also suggests that, in the absence of [ɲ] as a labeling category, the F2 difference of 250 Hz (1650-1900 Hz) between transition frequency values appropriate for [n] and [ɲ] (Table 3) was interpreted exclusively as [n] by English listeners (Table 4). On these grounds, it seems clear that stimuli with [ɲ]-like transitions were interpreted as [ŋ] by English listeners but might well have been categorized as [ɲ] by speakers of other languages, while stimuli with [ŋ]-like transitions, in the absence of an appropriate negative F3 and some [ŋ]-like murmur spectrum, were interpreted as [m] or [n]. This view is supported by observations reported in the literature. Thus experimenters at Haskins Laboratories stated that while, with no F3, "a large plus F2 transition is heard as [ɲ]... rather than [ŋ], with a -3 transition positive F2 transitions are now heard clearly as [ŋ]" (Haskins Laboratories QPR 8, 1953, pp. 21-22). Direct evidence about the effect of a strongly rising F2 in cueing palatals is also provided by Derkach, Fant, and Serpa-Leitão (1970) for Russian palatalized consonants. Thus, they found this F2 transition type to be the most relevant perceptual palatalization cue with vowel [a]. Consistently with the contrast in formant transition patterns for [ɲ] and [ŋ], important improvements in identification of the nonnasal stop [g] without F3 transition (Liberman et al., 1954) were found to be dependent on the presence of an optimal - 480 F3 (Harris, Hoffman, Liberman, Delattre, & Cooper, 1958). In studies on speech analysis listed in Table 3, the direction of F3 was found to be predominantly negative as well.

In the light of previous comments, there are reasons to believe that degrees of F1 excursion should be included as variable parameters in perceptual continua of place identification for nasals. Thus, as shown on other

Table 4

Frequency values (in Hz) corresponding to F2-F3 transitions in syllables with [a], [æ] and [m], [n], [ŋ] obtained from synthesis experiments with English. A summary of values is also included.

	Endpoints		Ranges		Direction		References
	F2	F3	F2	F3	F2	F3	
[m]	ar.700 (Locus)	—	—	—	-	-	(1)Liberman et al. (1954) ([a])
	900-1450	—	-190/ +360	—	-/+	—	(2)Nakata (1959)([a])
	ar.1300	—	-200	—	-	—	(3)Hecker (1962)([æ])
	1000-1300	2000-2300	-500	-200/ +100	-	-	(4)Garcia (1966, 1967a,b)([æ])
	920-1300	—	-310/ +70	—	-/+	—	(5)Miller & Eimas (1977)([a])
[n]	ar.1800 (Locus)	—	—	+360	+	+	(1)
	1450-1890	—	+360/ +800	—	+	—	(2)
	ar.180C	—	+300	—	+	—	(3)
	1300-1920	2400-3000	-500/ +120	+200/ +800	-/+	+	(4)
	1300-1700	—	+70/ +470	—	+	—	(5)
[ŋ]	ar.3000 (Locus)	—	—	-360	+	-	(1)
	1920-2300	—	+830/ +1210	—	+	—	(2)
	2000	—	+200	—	+	—	(3)
	1920-2300	2200-2400	+120/ +500	0/ +200	+	+/=	(4)

Summary

[m]	900-1450	-500/ +360	-200/ +100	-(+)	-(+)
[n]	1300-1890	-500/ +800	+200/ +800	+(-)	+
[ŋ]	1920-2300	+120/ +1210	-360/ +200	+	-(=, +)

grounds by Derkach et al. (1970), a strongly negative F1 transition is to be considered a relevant palatalization cue. The perceptual relevance of a higher F1 starting-ending point in identification of a non-nasal velar [g] vs. [d], [b] has been pointed out by Fujimura (1971).

Individual and overall duration of formant transitions can be shown to differ among place categories. It is significant that details about this aspect (longer transitions for [ŋ], shorter transitions for [m] and [n]) have been taken into consideration in works on speech synthesis (Hecker, 1962; Mattingly, 1968; Nakata, 1959). One must point out that such observations about relative transition duration for different categories mean little unless vocalic context is kept constant. In such circumstances, duration of F2 transition was found to be a rather important cue for the identification of different Polish nasal-vowel sequences (Kacprowski & Mikiel, 1965).

C. Interaction of Cues

In sections 2.2.A. and 2.2.B. I have taken into consideration how formant transitions and nasal murmurs contribute separately to the identification of nasal consonants of different place of articulation and what acoustic traits make transitions and murmurs perceptually relevant in such an identification process. Results show that:

(1) either [m] transitions or [m] murmur are sufficient place identification cues for [m]. The strong perceptual role of [m] murmur in [æ], [a] environments is to be particularly emphasized (Carlson, Granström, & Pauli, 1972; Henderson, Note 1; Malécot, 1956);

(2) [n] transitions are more powerful place cues than [n] murmur for [n] identification;

(3) [ŋ] cues, transitions, murmur, and release, are needed for a satisfactory [ŋ] identification with vowels [æ], [ɛ] but not with [a]. In this case [ŋ] murmur is a stronger cue than [ŋ] transitions;

(4) [ɲ] transitions but not [ɲ] murmurs are sufficient place identification cues for [ɲ].

The arrangement of cues in running speech suggests that place cues for nasals ought to be explored interdependently instead of in an isolated way. Any attempt to detect them should focus primarily on the interactive role of transitions, murmurs, and releases. Results reported by experiments with natural and synthetic speech, already discussed, can be said to adduce some valuable (although indirect) information to this issue. Such experiments presented murmurs in isolation for identification (Dukiewicz, 1967; House, 1957; Malécot, 1956; Manrique de et al., 1980; Nakata, 1959) or murmurs or transitions combined with the vocalic steady state portion, with or without final release (Malécot, 1956; Henderson, Note 1). But, none of such attempts succeeded in combining all cues reciprocally to detect possible cross-category effects. A more realistic approach is reported by Malécot (1956) who, in addition to the experiments mentioned above, compared the cue value of transitions and released murmurs extracted from natural speech utterances with [æ] and [m], [n], [ŋ]; stimuli were tested with American English subjects.

According to Malécot's data, [m] cues (murmur and transitions) override transitions and murmurs for [n] and [ŋ], and [n] cues (murmur and transitions) override transitions and murmur for [ŋ]. These overriding effects are significant in all cases except for [n] murmur upon [ŋ] transition (52% to 48%). They report valuable information about cross-category effects of transitions vs. murmurs and, therefore, are of particular interest for my experiment on Catalan nasals.

D. Summary

It has been shown that place cues for nasals are complex. In order to characterize them satisfactorily, experiments with natural and synthetic speech need to be appropriately designed. Some relevant suggestions about this subject, to be taken into consideration in further research and partly accounted for in my perceptual experiment on Catalan nasal consonants, are presented in the following paragraphs.

Synthetic stimuli have to recreate the temporal arrangement of cues found in natural speech. Results about perceptual relevance of cues summarized in the preceding section are to be taken with care. Thus, except for those reported in Malécot's experiment on interaction of cues, they derive from experimental paradigms in which the arrangement of cues in the stimuli can be shown to deviate from the arrangement of cues observed in the syllabic structure. This is not only the case for isolated murmurs but also for combinations of segments: abutting the remaining portions of the signal (Malécot, 1956; Henderson, Note 1) when murmurs or transitions have been removed clearly alters the timing relationship between all VC cues and, as Henderson herself has pointed out, might cause masking phenomena when transitions and final release are presented in succession; on the other hand, preserving the timing relationship between remaining cues would leave unnatural silent portions in the stimuli.

Malécot's experiment on interaction of cues accounts for their temporal arrangement in the syllable but does not provide any evidence for place information conveyed by different transition characteristics and murmur spectral regions. A speech synthesis paradigm is needed for this purpose in which frequency values are close to analysis data of real speech. Moreover, acoustic traits of transitions and murmurs need to be varied simultaneously so that results do not refer independently to optimal transition cues and optimal murmur cues but to unitary transition-murmur cues. In fact, data reported in this paper reveal that murmurs and transitions for [m], [n], [ɲ], [ŋ] are perceptually complementary similarly to bursts and transitions for non-nasal stops and, consequently, are integrated analogously in the perceptual process with reference to well-stated production constraints.

In the following experiment, I try to put into practice these views on research strategies. I investigate the perceptual effects of acoustic cues in combination, using Catalan subjects, by contraposing transition and unreleased murmur patterns for final nasal consonants of different place of articulation. In contrast with the experimental paradigms reviewed, the arrangement of cues within the syllabic structure is preserved while varying transition and murmur simultaneously and systematically. Although the perceptual relevance of all individual acoustic parameters is not tested, such a dynamic approach allows a better understanding of their role within the overall transition and murmur patterns as well as the VC structure. Analysis data that have been reviewed and theoretical issues that have been raised as well as additional information obtained from Catalan speakers are taken into consideration.

3. AN EXPERIMENTAL STUDY ON CATALAN NASAL CONSONANTS

1. Phonetic and Phonological Description

In Eastern Catalan, phones [m], [n], [ɲ], [ŋ] appear in absolute final position. [m], [n], [ɲ] also appear intervocalically and correspond to underlying /m/, /n/, /ɲ/. [ŋ] is found word-internally only immediately before [k] and, in final position, occasionally in free variation with [ŋk] depending on the speaker and the lexical item. These phonetic facts argue for [ŋ] being an allophone of underlying /n/ before a velar stop, generated by the following phonological derivation (see also Mascaró, 1978):

"Underlying form"	/sang/	/bank/
Regressive Assimilation	ŋ	ŋ
Optional Deletion	(0)	(0)
Devoicing	(k)	
"Surface form"	[saŋ(k)] 'blood'	[baŋ(k)] 'bench'

The presence of underlying /g/ and /k/ accords well with the phonetic realization of derived formations such as [səŋgi'nari] 'bloodthirsty', [səŋgu'nos] 'bloody', [bəŋ'kɛtə] 'low bench'. Other minimal pairs with contrasting nasal stops in final position can be found. Thus, [fam] 'hunger', [fan] '(they) do', [faŋ] 'mud'; [bam] 'Aux. we go', [ban] 'edict', [baɲ] 'bath' [baŋ] 'bench'.

Final [m], [n], [ɲ], [ŋ] are weakly released or unreleased according to individual speakers. Given the occurrence of the full set of nasals of different place of articulation and the unreleased murmur condition, it is possible, then, in Eastern Catalan, to investigate the role that transitions and unreleased murmurs play in the process of identification of different final place categories.

2. Production Study

Analysis values corresponding to production data from a single male Catalan speaker were chosen as reference points for the patterns to be used in synthetic speech experiments. Several samples of monosyllabic minimal pairs were analyzed by means of a Voice Identification sound spectrograph, a digital spectrographic analyzer and a linear prediction model analysis. To see to what extent acoustic patterns found in production data on nasals from this reference speaker were in accordance with those of other Catalan informants, data on the same minimal pairs embedded in a neutral sentence were also collected from 12 other male Catalan speakers and analyzed spectrographically. Frequency values for the reference speaker as well as range of frequency variation and mean frequency values for the other 12 speakers are presented in Tables 5 and 6. Both sets of data are compared below and discussed in the light of theoretical predictions and data from the literature.

Nasal murmur values for the reference Catalan speaker (Table 5) are consistent with those given in Table 1. The structural configuration of poles and zeros in Table 1 is violated by N3 and N4 for [ɲ], which, as for the other Catalan speakers, happen to be higher than that for [ŋ].[5] The continuity between F3 transition and N4 for [ɲ] suggests that palatal N4 is mouth cavity dependent. Large N1 bandwidth values for [ŋ] are consistent with data obtained by other investigators (section 2.2.A.).[6] As shown in the same table, the other Catalan speakers differ from the reference speaker in that they fail to show a contrast between [n] and [ŋ] with respect to N3 and N4. This fact suggests that, if murmur spectra are found to convey relevant information in discriminating [n] from [ŋ], perceptual distinctiveness is to be assigned to contrasting values for N1 and NZ. It would be interesting to test the index of perceptual confusability between [n] and [ŋ] in Catalan using real speech stimuli with tokens of unreleased [ŋ].

For Catalan speakers (the reference speaker and 12 others) the general direction corresponding to F2-F3 transition mean values (Table 6) is consistent with that reported in the current literature on synthetic speech experiments for [m], [n], [ŋ] in English (Table 4). Also cross-category F1 values can be predicted on the basis of the acoustic theory of speech production. F2, not only for [ɲ] but also for [n], is consistently positive (see Table 3 for comparison), even when extremes of the observed range of values are taken into consideration. While F2 values for [m] and [ɲ] fall well apart from those for [n] and [ŋ], respectively, values for [ŋ] overlap significantly with those for [n] and even [m]. This is consistent with the fact that an appropriate F3 is needed to synthesize a satisfactory, unambiguous velar nasal consonant.

The nasal murmur was at least 1.5 times longer than the preceding vocalic segment for the reference Catalan speaker. Transition durations for [ɲ] were consistently longer (70 msec average) than those for [n] (50 msec average) and [ŋ] (35 msec). For this speaker, as well as for many other Catalan speakers, positive F2-F3 excursions were still observable during the nasal murmur period as an effect of the dynamic motion exhibited by the large mass of tongue body towards the dorsopalatal region. The perceptual relevance of the timing relationship between the gliding movement and the nasal closure onset has not been investigated in my perception experiment: nasal formants were kept steady instead, as found in the productions of some Catalan speakers. Murmur release and final voiceless stop after [ŋ] were present or absent as predicted in section 3.1.

3. Perceptual Study

A. Procedure

To explore the perceptual role of transitions and unreleased nasal murmurs in place recognition as well as to detect identification cues for [ɲ], continua with [an], [aɲ], [aŋ] were synthesized in two parallel blocks of two slightly different tests each (1a, 1b; 2a, 2b) according to analysis values

Table 5

Analysis frequency values (ranges and means) (in Hz) for murmurs in VC syllables [am], [an], [aɲ], [aŋ] according to data from a single Catalan speaker and 12 other Catalan subjects.

	N1	N2	N3	N4	NZ	Subjects
[m]	200	1120	1360	2100	—	(1)Single subject
	170-320	910-1105	1120-1510	1370-1800	—	(2)12 subjects (range values)
	255	1015	1300	1565	—	(3)12 subjects (mean values)
[n]	200-300 BW:180	800-900	1460-1650 BW:330	1950-2100 BW:235	1780	(1)
	225-325	880-1135	1440-1640	1775-2600	—	(2)
	285	1035	1515	2130	—	(3)
[ɲ]	180-230 BW:150	900-1150 BW:220	2000-2200 BW:140	2900-3350 BW:140	2360-3000	(1)
	200-340	800-1180	1365-2335	1740-3000	—	(2)
	295	1055	1760	2265	—	(3)
[ŋ]	300-400 BW:275	1150-1240 BW:200	1860-2200 BW:300	2430-2650 BW:250	2900-3400	(1)
	225-360	900-1240	1375-1640	1960-2730	—	(2)
	295	1060	1530	2160	—	(3)

Table 6

Analysis frequency values (ranges and means) (in Hz) for transitions in VC syllables [am], [an], [aɲ], [aŋ] according to data from a single Catalan speaker and 12 other Catalan subjects

	V steady-state			Endpoints			Subjects
	F1	F2	F3	F1	F2	F3	
[m]	920	1435	2420	800	1150	2420	(1)Single subject
	630-	1110-	1940-	525-	980-	1880-	(2)12 subjects
	825	1580	2250	765	1560	2250	(range values)
	725	1290	2130	655	1215	2075	(3)12 subjects
							(mean values)
[n]	920	1535	2520	775	1540-	2600	(1)
					1630		
	570-	1200-	1955-	525-	1350-	1955-	(2)
	805	1510	2610	760	1660	2610	
	715	1325	2300	685	1560	2365	(3)
[ɲ]	920	1535	2520	450-	1840-	2750-	(1)
				550	1980	2850	
	685-	1200-	1975-	405-	1625-	2025-	(2)
	875	1570	2700	805	2050	2925	
	755	1400	2315	635	1860	2440	(3)
[ŋ]	920	1535	2520	900	1400-	2125-	(1)
					1640	2500	
	685-	1285-	1995-	645-	1310-	1860-	(2)
	900	1680	2700	900	2025	2475	
	770	1440	2255	740	1550	2140	(3)

	Ranges			Direction			
	F1	F2	F3	F1	F2	F3	
[m]	-120	-285	0	-	-	=	(1)
	-155/	-270/	-110/	-	-(=,	-(=)	(2)
	-25	+30	0		+)		
	-70	-75	-55	-	-	-	(3)
[n]	-145	-5/	+80	-	+(=)	+	(1)
		+95					
	-90/	+115/	-190/	-(=)	+	+(-, =)	(2)
	0	+300	+285				
	-30	+235	+65	-	+	+	(3)
[ɲ]	-470/	+305/	+230/	-	+	+	(1)
	-370	+445	+330				
	-425/	+200/	-285/	-(+, =)	+	+(-, =)	(2)
	+50	+630	+435				
	-120	+460	+125	-	+	+	(3)
[ŋ]	-20	-135/	-395/	-	-/+/=	-	(1)
		+105	-20				
	-100/	-75/	-315/	-(=)	+(=)	-(+, =)	(2)
	+10	+450	+90				
	-30	+110	-65	-	+	-	(3)

for the Catalan reference speaker displayed in Tables 5 and 6. The synthetic stimuli were prepared using a software serial formant synthesizer (SYNTH) at Haskins Laboratories with variable BW parameters, an extra pole (used as N2) and an extra zero (used as NZ) (Mattingly, Pollock, Levas, Scully, & Levitt, 1981).

In tests 1a and 1b a series of variable F2-F3 transition endpoints was combined with three fixed murmur patterns believed to be optimal for [n], [ɲ], [ŋ]; in tests 2a and 2b a series of variable murmur values was combined with three fixed, optimal transition patterns. In contrast with previous experimental studies, these two conditions allow us to determine identification frequency ranges across place categories for transitions and murmurs as well as to investigate more adequately the interaction of the two acoustic cues within a syllable structure recreated from that of natural speech utterances. Actual values are given in Figures 1 and 2. Poles of the murmur pattern are represented by single lines and zeros by double lines. Test 1a differs from 1b and 2a from 2b in vowel steady state and transition endpoint values. Two versions of the same experimental design were given simply to test the perceptual effect of a larger variety of F2-F3 transition values and, therefore, to be able to determine identification cross-over points with higher precision. I give some details about the variable transition endpoints (Figure 1) and variable murmur structures (Figure 2), which will be taken into consideration in the discussion section on perceptual data obtained from Catalan informants:

a. F2 transition endpoints:

Test 1a: From 1430 Hz (-70) to 1920 Hz (+420) in 7 steps of 70 Hz;
Test 1b: From 1600 Hz (0) to 2000 Hz (+400) in 8 steps of 50 Hz.

b. F3 transition endpoints:

Test 1a: From 2340 Hz (-160) to 2900 Hz (+400) in 7 steps of 80 Hz;
Test 1b: From 2300 Hz (-300) to 3100 Hz (+500) in 8 steps of 100 Hz.

c. Nasal murmur structures:

N1: From 250 Hz (6 steps) to 350 Hz in 5 steps of 20 Hz;
N2: From 900 Hz to 1200 Hz in 11 steps of 30 Hz;
N3: From 1600 Hz to 2100 Hz (6 steps) in 5 steps of 100 Hz;
N4: From 2500 Hz to 3000 Hz and vice versa in 5 steps of 100 Hz;
NZ: From 1800 Hz to 3200 Hz in 11 steps of 140 Hz;

Formant bandwidth values were also varied across stimuli according to frequencies included in Figures 1 and 2. A constant value of 900 Hz for F1 was chosen. Preceding [a] was nasalized from vowel onset to transition endpoint by means of a progressive frequency rise of a single low pole (500 to 600 Hz) - zero (500 to 700 Hz) pair. Each stimulus was 560 msec long, having a vowel steady state of 200 msec, a transition of 60 msec and a murmur portion of 300 msec. There was a progressive 10 dB decrease from vowel to murmur and an F0 lowering slope between vowel onset (120 Hz) and murmur offset (80 Hz).

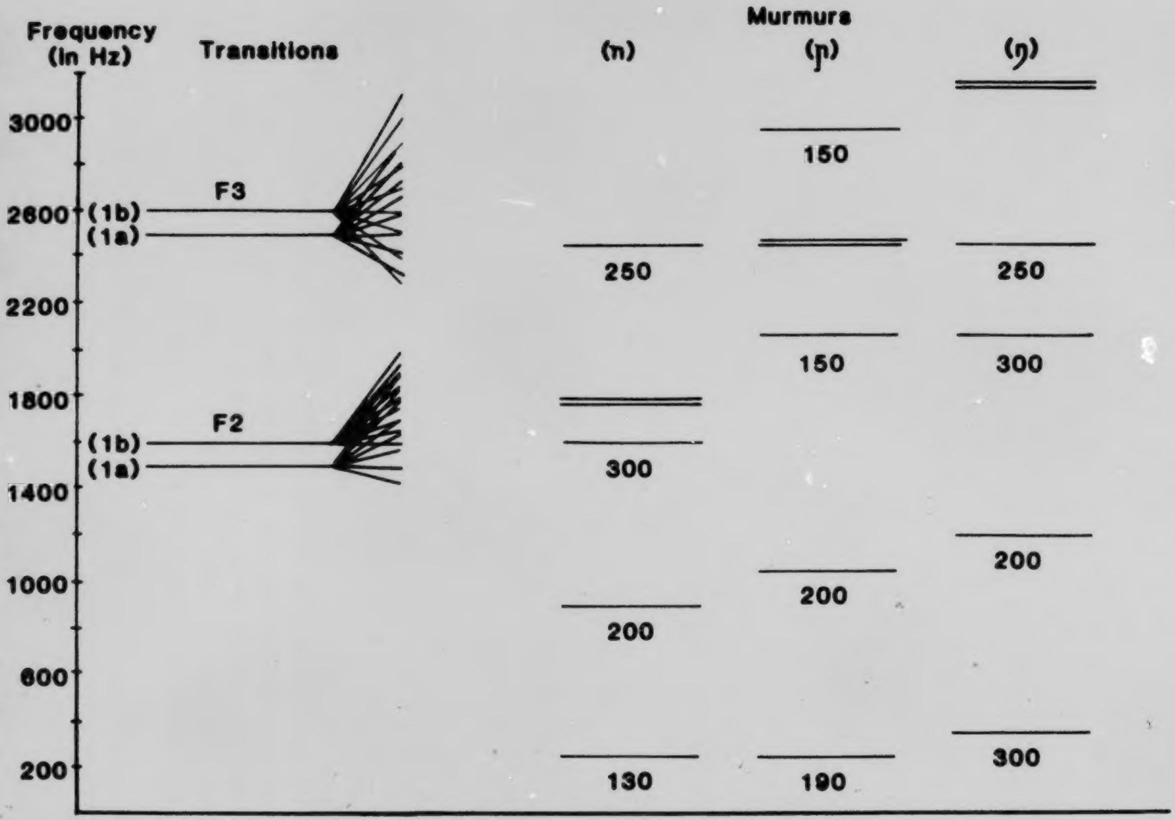


Figure 1.- Synthesis patterns for tests 1a and 1b (fixed murmur conditions) with inclusion of bandwidth values (in Hz) for murmur formants.

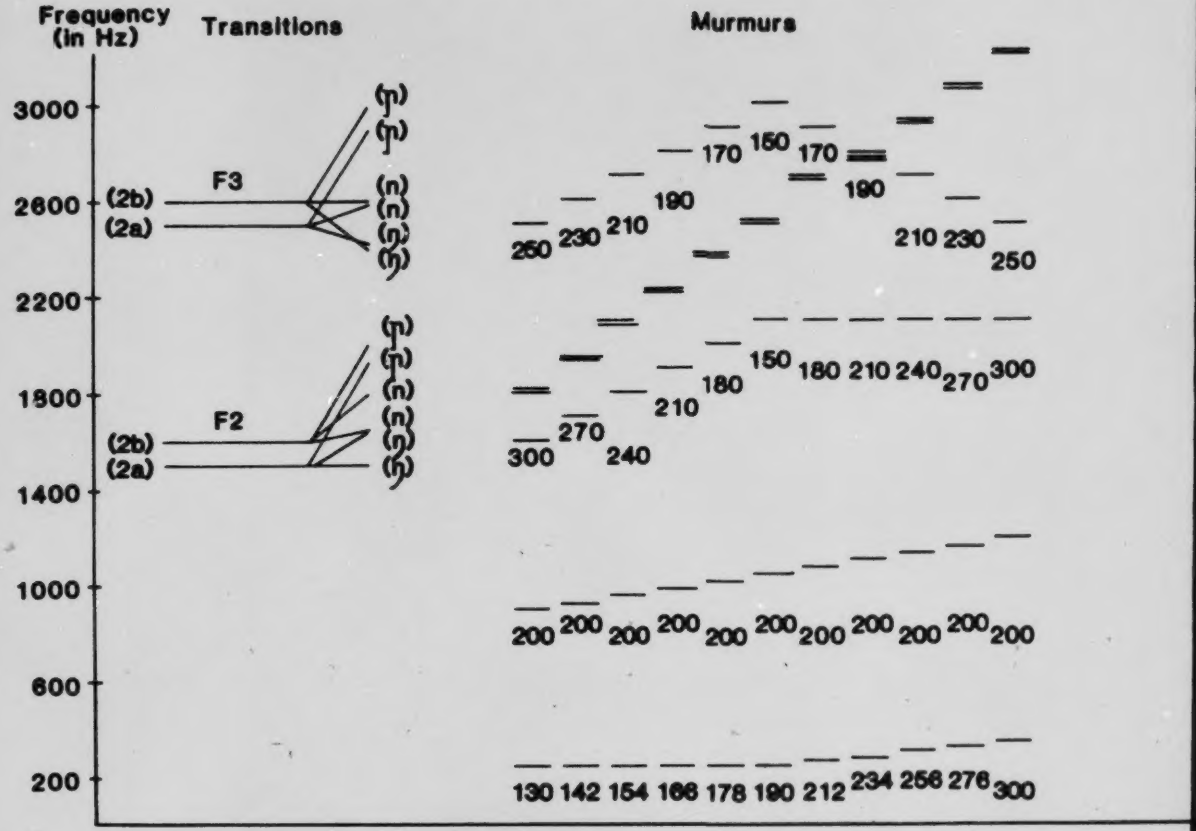


Figure 2.-Synthesis patterns for tests 2a and 2b (fixed transition condition) with inclusion of bandwidth values (in Hz) for murmur formants.

Every test was administered using several stimuli per step and the overall set of stimuli randomized before presentation for identification. Overall, test 1a was composed of 144 tokens (6 per step), test 1b of 162 (6 per step), test 2a of 165 (5 per step) and test 2b of 176 (4 per step). Intervals of 4 sec were included between successive stimuli and longer 10 msec intervals after every ten stimuli. Twenty-four paid Catalan subjects took each of the four tests twice and were asked to identify orthographically the final nasal stop as [n], [ɲ] or [ŋ]. They were all students who had had no previous experience with synthetic speech and did not know the purpose of the experiment. Thirteen took the tests binaurally through headphones; the rest listened to stimuli reproduced through a loudspeaker because of problems connected with taking the tests in the field. To find out whether such different listening conditions could have had some strong effect on subjects' responses, I listened to stimuli under both conditions and obtained almost identical cross-over points and response distributions.

B. Results

Table 7 shows category judgments for each test in all variable conditions; data from all subjects and from the most consistent 14 labelers are reproduced separately. Results for each variable condition have been displayed in Figures 3, 4, 5, and 6. Figures 3 and 4 give perceptual data obtained from tests 1a and 1b, and Figures 5 and 6 give data from tests 2a and 2b. Figures 3 and 5 give data from all 24 Catalan subjects and in Figures 4 and 6 data from the most consistent labelers (14 selected subjects). Each subplot in Figures 3 and 4 represents judgments for a particular murmur; stimulus numbers for different F2-F3 transition values lie on the abscissa. Each subplot in Figures 5 and 6 represents judgments for a particular F2-F3 transition set; stimulus numbers for different murmur values lie on the abscissa. Among all subjects those who identified the stimuli with most consistency were chosen as "best" labelers. A comparison of identification curves obtained from these 14 selected informants with those obtained from all 24 shows that, as expected, the best labelers categorized stimuli more distinctively. Thus, the following summary on perceptual data about [n], [ɲ], [ŋ] identification will refer mainly to responses obtained from the most consistent Catalan subjects.

a. Table 6.-A comparison of percentages of category identification between murmurs and transitions shows that the labeling of appropriate transitions is always more consistent than that of appropriate murmurs. The effect of transitions vs. murmurs in both sets of tests (1a-2a, 1b-2b) is much higher for [ɲ] (2.5-2.8 ratio) than for [n], [ŋ], and for [n] (1.4-1.7 ratio) than for [ŋ] (1.1-1.6 ratio).

Tests 1a and 1b show that an optimal [ɲ] murmur does not favor the identification of one or another of the place categories. An optimal [n] murmur slightly favors identification of [n], [ɲ] vs. [ŋ] (see test 1b). An optimal [ŋ] murmur favors significantly [ŋ] vs. [n] responses (1.5-1.9 ratio).

Tests 2a and 2b show that the presence of optimal [ŋ] transitions correlates significantly with [ŋ] vs. [n] responses (1.1-2.1 ratio) while that of optimal [n] transitions correlates significantly with [n] vs. [ŋ] responses (1.1-2.1 ratio), independently of [ɲ] identification. Appropriate [ɲ] transitions contribute exclusively to [ɲ] identification in a range of 90.9%-95.7%.

Table 7

Category judgments for each test (percentage). Data from all 24 subjects and the best 14 labelers are displayed separately.

	[ŋ]	[n]	[ɲ]	Subjects
Test 1a				
[ŋ] murmur	40.8	26.6	32.6	(1)All 24 subjects
	42.7	24.9	32.2	(2)Best 14 labellers
[n] murmur	30.3	32.9	36.8	(1)
	28.4	34.5	37	(2)
[ɲ] murmur	33	31.2	35.8	(1)
	32.3	34.4	33.3	(2)
Test 1b				
[ŋ] murmur	44.1	23.6	32.3	(1)
	45.1	22.6	32.2	(2)
[n] murmur	26.7	37.1	36.1	(1)
	23.3	40.4	36.3	(2)
[ɲ] murmur	32.5	31.5	36	(1)
	32.3	34.7	33	(2)
Test 2a				
[ŋ] transtions	63.8	35.3	0.4	(1)
	68.4	31.1	0.5	(2)
[n] transitions	46	51.2	2.8	(1)
	39	60.4	0.4	(2)
[ɲ] transitions	1.5	2.8	95.7	(1)
	0.6	4.3	95.1	(2)
Test 2b				
[ŋ] transitions	53	45.4	1.6	(1)
	54	45.9	0.2	(2)
[n] transitions	37.9	54.8	7.1	(1)
	30.4	65.7	3.9	(2)
[ɲ] transitions	2.8	4.1	93.1	(1)
	2.6	6.5	90.9	(2)

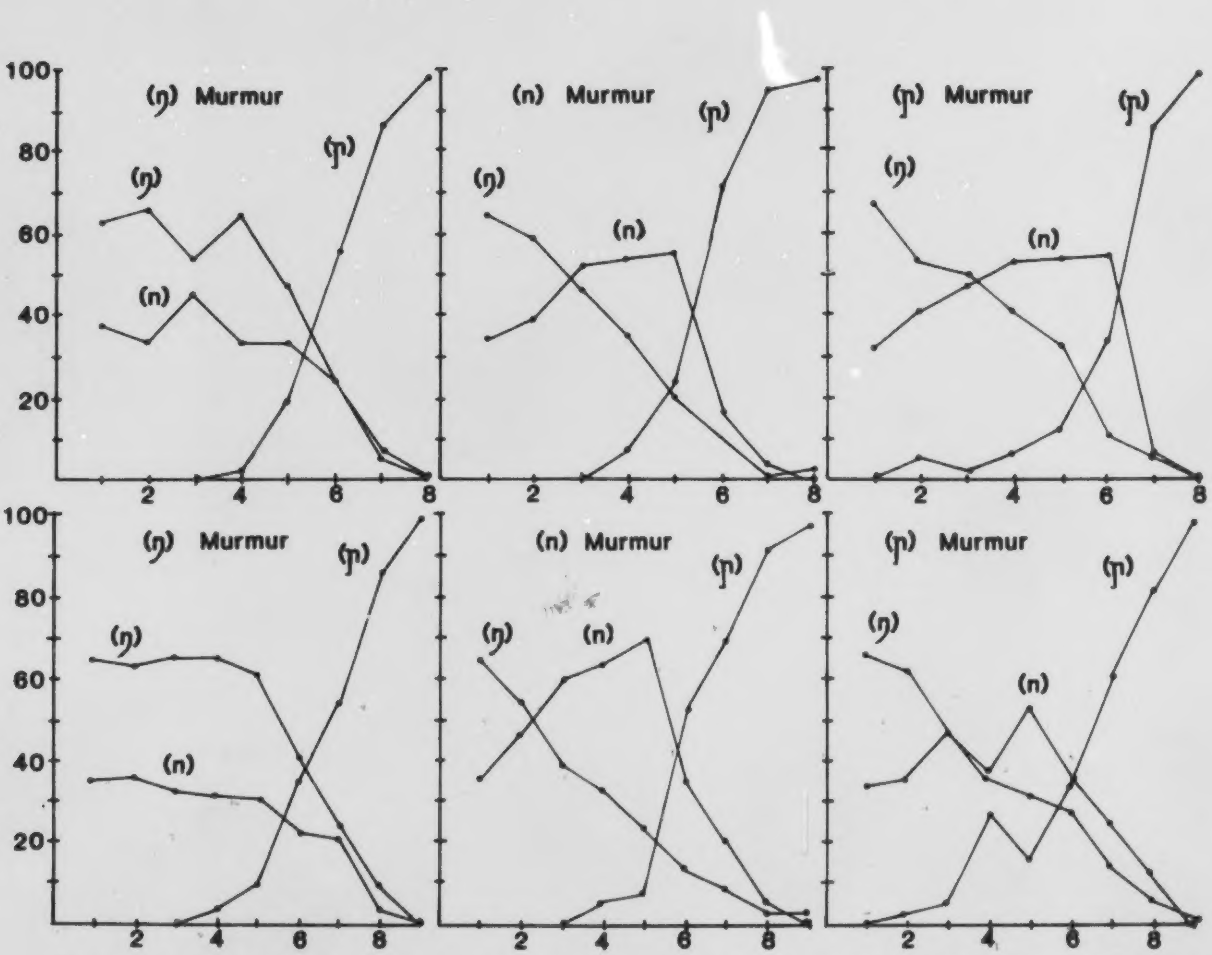


Figure 3.- Perceptual results for tests 1a (above) and 1b (below) (all 24 Catalan subjects). Ordinate:percentage of responses with indicated label; abscissa:stimuli.

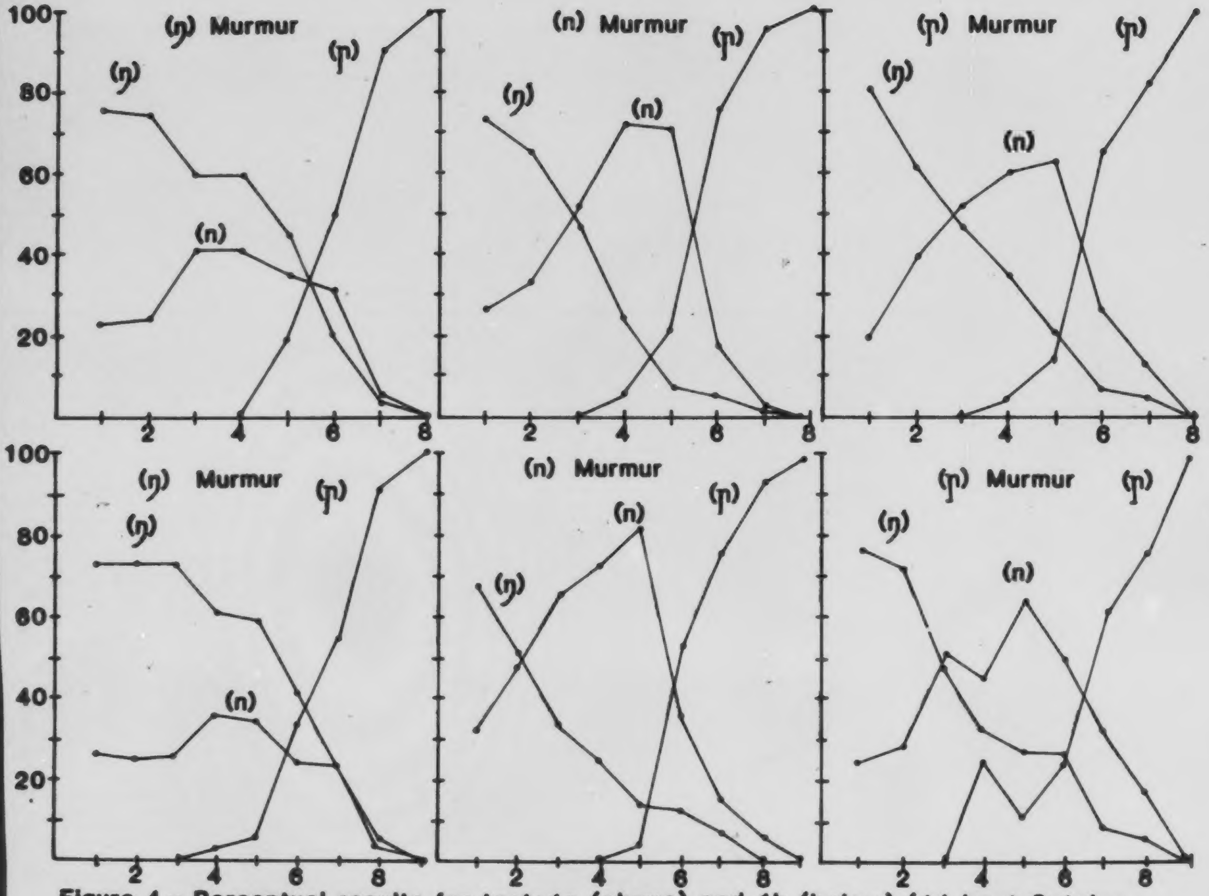


Figure 4.- Perceptual results for tests 1a (above) and 1b (below) (14 best Catalan labelers). Ordinate: percentage of responses with indicated label; abscissa: stimuli.

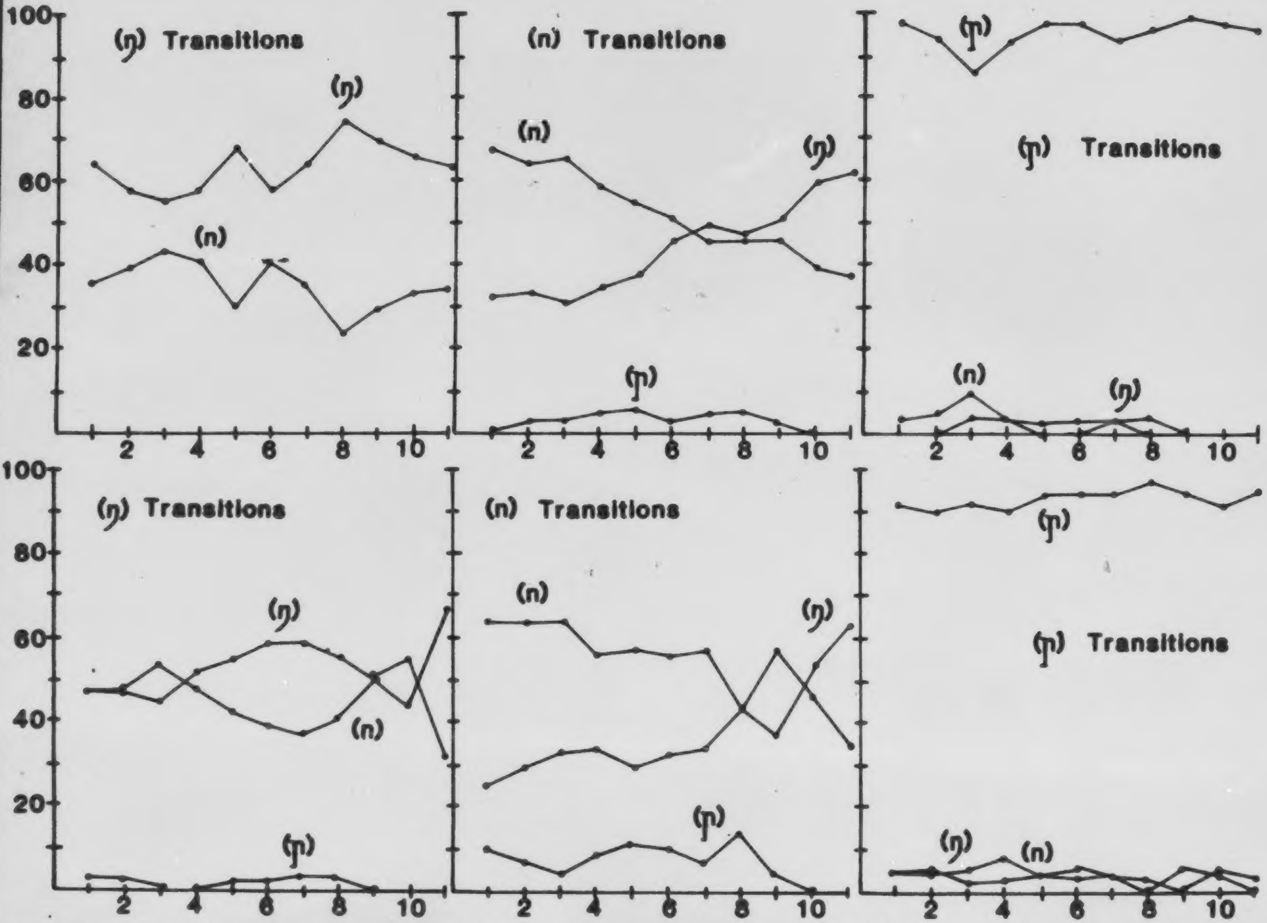


Figure 5.–Perceptual results for tests 2a(above)and 2b(below)(all 24 Catalan subjects). Ordinate:percentage of responses with indicated label; abscissa:stimuli.

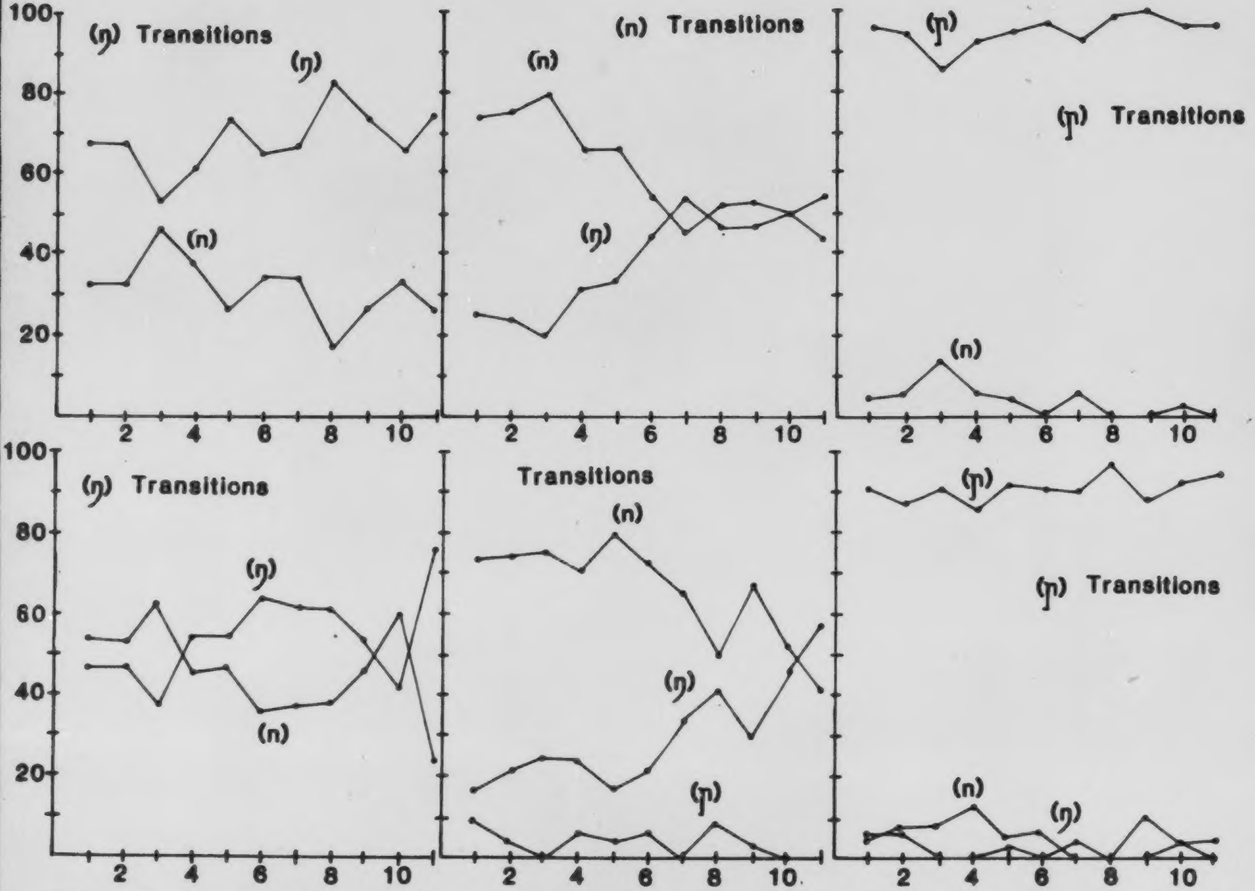


Figure 6.-Perceptual results for tests 2a (above) and 2b (below) (14 best Catalan labelers). Ordinate: percentage of responses with indicated label ; abscissa: stimuli.

b. Tests 1a, 1b (Figures 4 and 5)-Identification peaks show that category judgments cannot be predicted on the basis of appropriate murmurs but of appropriate transitions, especially for the [n] and [ɲ] murmur conditions. I characterize below an optimal set of F2-F3 transition directions and range values (in Hz) for each place category, according to perceptual data reported by all Catalan speakers with special reference to the most consistent labelers:

[ŋ]	F2: slighty negative to slightly positive	(-70 to +80)
	F3: strongly negative to steady	(-300 to 0)
[n]	F2: positive	(+140 to +250)
	F3: steady to positive	(0 to +200)
[ɲ]	F2: strongly positive	(+280 to +420)
	F3: strongly positive	(+200 to +500)

Murmurs appropriate for different categories have no effect upon optimal transition values for [ŋ] and [ɲ]. While [ɲ] murmur has no effect upon optimal [n] transitions, a significantly higher average of [ŋ] than [n] judgments is obtained for typical [n] formant values (F2: up to +260; F3: up to +230) followed by [ŋ] murmur.

b. Tests 2a, 2b (Figures 5 and 6)-Identification peaks show that category judgments can be highly predicted on the basis of appropriate transitions. This is clearly the case for the [ɲ] transition condition: while no optimal [ɲ] murmur can be found along different murmur continua, optimal [ɲ] transitions override completely [n] and [ŋ] murmurs. Therefore, an optimal set of murmur values (in Hz) for [ŋ] and [n] is to be exclusively found in the case of the [ŋ] and [n] transition conditions:

[ŋ]	N1: 350	N2: 1200	N3: 2100	N4: 2500	NZ: 3200
[n]	N1: 250	N2: 960	N3: 1800	N4: 2700	NZ: 2080

The perceptual effect of these optimal murmur values is obvious: for [ŋ] and [n] transitions the percentage of [ŋ] responses increases as the optimal [ŋ] murmur (stimulus 11) is approached and that of [n] responses also rises towards the optimal [n] murmur (stimulus 3). It is the case that optimal murmurs for [ŋ] and [n] are shown to be dependent upon [n] and [ŋ] transitions respectively: thus, the percentage of [ŋ] responses for the optimal [ŋ] murmur is significantly lower with [n] than [ŋ] transitions and that of [n] responses for the optimal [n] murmur is significantly lower with [ŋ] than [n] transitions. Moreover, it must be noticed that, while [n] transitions never override optimal [ŋ] murmurs, [ŋ] transitions are shown to override optimal [n] murmurs in test 2a (Figure 6).

C. Summary and Discussion

Perceptual data from Catalan subjects indicate that, overall, for vowel [a], transitions provide more effective cues for nasal consonants of different place of articulation than murmurs. This is consistent with results obtained in previous experiments with synthetic speech tested with American English speakers. In agreement with Henderson's (Note 1) results, it has been shown, however, that the contribution of murmurs in place identification is much higher than that suggested by data reported in most of those experiments, and category-dependent as well. In fact, my results confirm that a better characterization of place cues for nasals must be strongly related to a considerable improvement in the construction of the experimental paradigm selected for perceptual testing. I draw next a summary of the perceptual results on Catalan and evaluate them in the light of material reviewed in Section 2.2.

While, as stated, transitions are more powerful cues than murmurs in the process of identification of nasals, important cross-category differences can be established. This effect is highly relevant for [ɲ], and more relevant for [n] than [ŋ]. Reciprocally, [ŋ] murmurs contribute more to place identification than other murmurs. This contrast in perceptual relevance of cues is consistent with the tendency of [ŋ] murmur to prevail over [n] transitions and not vice versa in what may be called an inter-category trading relation as characterized below in this discussion section. On the other hand, while [ɲ] transitions override murmurs appropriate for other nasal categories, optimal [ɲ] murmurs have been shown to convey no place information. This negative effect may have been maximized by not having taken into consideration, in the synthetic reproduction of murmurs, the characteristic [ɲ] glide component during murmur. Spectrographic analysis reveals that, while very little movement can be detected for nasal formants during the closure period in the case of [n] and [ŋ], those for [ɲ] show a continuation of positive excursion with respect to F2-F3 transitions. Since F-transitions result from articulatory dynamics, we have, apparently, a continuation of tongue movement during lingual closure and complete oronasal coupling. All these findings about relevance of transitions and murmurs in the identification of different place categories are significantly consistent with the summary of interactive perceptual effects included in section 2.2.C., based upon production and perception data from other languages.

Optimal cues obtained for [ŋ] indicate that transition direction for F2 is not perceptually relevant as long as it remains close to the vowel steady state frequency; however, F3 transition must be negative for a satisfactory [ŋ] identification. Optimal [ŋ] murmur, on the other hand, is characterized by a high N1 and the absence of NZ at the central part of the spectrum. Identification of [n] is mainly dependent upon a constantly positive F2, for a steady or positive F3. Optimal alveolar murmurs have a low N1 and an NZ between N3 and N4. These results for [ŋ] and [n] conform well with indications about perceptually relevant acoustic cues in sections 2.2.A. and 2.2.B. according to data from other languages. [ɲ] has been found to be exclusively dependent on strongly positive F2-F3 transitions in agreement with reported perceptual data from Polish and cues for Russian palatalized non-nasal consonants. This powerful transition effect also confirms suggestions made in section 2.2.B. about the possibility of a [ɲ] identification for a strongly positive F2 transition in the absence of appropriate [ŋ] cues by speakers of languages with [ɲ]. Moreover, the fact that no perceptual effect follows the contrast in N3-N4 values between [ɲ] murmur and [n], [ŋ] murmurs (Figure 1) is consistent with the irrelevance of high formants at the murmur portion in place identification of nasal murmurs and, therefore, with the perceptual significance of N1 and NZ values. In this summary of cues one must point out that the inclusion in the experimental paradigm of different F1 transitions for contrasting nasal categories might have added some relevant information about place identification.

Data in Figure 7 show to what extent the perceptual results are consistent with production measurements gathered from Catalan speakers. Crosses along both diagonal lines point to values for F2-F3 transition ranges corresponding to synthetic stimuli of tests 1a and 1b. Phonetic symbols recorded on these lines indicate prevailing interstimuli category judgments for transition continua in all different murmur conditions. Dots stand for F2XF3 transition range values corresponding to productions of [m], [n], [ɲ], [ŋ] by single Catalan speakers summarized in Table 6; they are grouped into production spaces for each nasal category. Transition range values for the reference Catalan speaker chosen to prepare the synthetic speech stimuli are represented by encircled dots. Such a graph has been found to accord more with perceptual processing of nasals than one that relates stimulus points to production data on transition endpoint values: A comparison of results from tests 1a and 1b has shown that, in categorizing stimuli, listeners were in fact attending to F2-F3 transition ranges relative to the vowel steady state value and not to absolute transition endpoints. A satisfactory coincidence is found between perceptual category judgments and category production spaces for different nasals, thus confirming the fact that transitions are good identification place cues. While this is clearly the case for the perceptual contrast between [n] and [ɲ],[7] it can be seen that murmur structures ([ŋ] murmur for [ŋ] identification; [n], [ɲ] murmurs for [n] identification) are used by Catalan speakers as identification cues for F2-F3 range values that lie somewhere between or on the edges of [ŋ] and [n] production spaces. This finding argues for the existence of a trading relation between acoustic traits spread over time in the process of [ŋ] vs. [n] identification, and shows that acoustic cues are integrated into a unitary phonetic percept in the process of dynamic perception. Thus, perceptual complementarity between transitions and murmurs accords with the fact that, given an ambiguous set of F2-F3 formant transitions between [ŋ] and [n] production spaces, listeners report /ŋ/ or /n/ judgments depending on whether the following murmur structure is appropriate for [ŋ] or [n], respectively. Moreover, consistent with reported observations, [ŋ] murmur appears to have greater perceptual weight than [n] murmur, since the perceptual range for [ŋ] affects more considerably the [n] production space than that for [n] the [ŋ] production space.

4. CONCLUSIONS

In the previous sections I have investigated the interactive role of formant transitions and unreleased murmurs in the process of identification of [n], [ɲ], [ŋ] with [a] in VC syllables using synthetic speech stimuli. Perceptual results from Catalan speakers strongly suggest that syllabic cues such as transitions and murmurs are simultaneously processed in a phonetic mode. As Studdert-Kennedy (1977) has made clear, dynamic acoustic events "are jointly shaped by the timing mechanisms of motor control and by the demands of the auditory system for perceptual contrast and compression" (p.17). As exemplified below, production and perception data reported in this paper show that there is evidence for both related strategies in the process of identification of place for nasals, namely, reference to articulatory constraints and to constraints imposed by the auditory system.

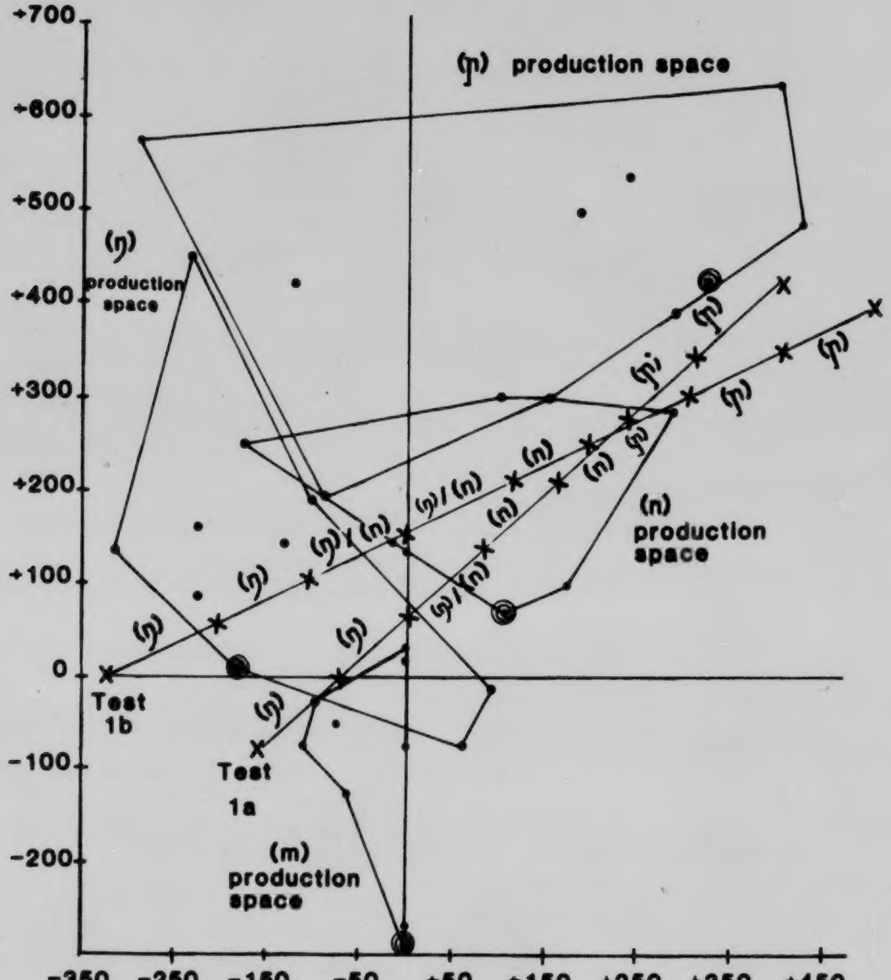


Figure 7.- Comparison of perception (Figures 4,6) and production (TABLE 6) data for Catalan subjects with special reference to formant transitions.Ordinate: F2 ranges (in Hz);abscissa: F3 ranges (in Hz).

Reference to psychoacoustic constraints imposed by the auditory system is presumably needed to account for specific acoustic cues that make transitions and murmurs perceptually relevant. On the one hand, it would argue for the correlation between amount of F2 transition range and perceptual relevance of F2 transition found to be true for nasals with respect to the arrangement [ɲ]> [n]> [ŋ]. An explanation for this effect is suggested by Klatt and Shattuck (1975). They found in experiments with non-speech stimuli that the perceptual importance of an F2-like chirp with respect to an F3-like chirp is positively correlated with its frequency height. That is, the effect would increase with an F2 transition such as that for [ɲ] (strongly positive) with respect to [n], [ŋ] and such as that for [n] (moderately positive) with respect to [ŋ] (slightly positive but also slightly negative or steady). No auditory constraint is known that can handle perceptually distinctive cues of [ŋ] vs. [n] murmurs.

An auditory analysis of that sort is compatible with a perceptual processing mechanism of relevant acoustic cues that keeps track of the underlying unitary articulatory gesture. In fact, it is to be thought that constraints imposed by the auditory mechanism in the interpretation of cues are integrated at a more central stage with reference to a dynamic and continuous set of coarticulation strategies. Evidence for such related events can be derived from Figure 7: according to data displayed there, nasals are perceived with reference to well-established category spaces and essentially processed, at least for [a], upon F2-F3 transition range values and upon murmur characteristics for potentially "ambiguous" transition configurations. Reference to the continuous production event is also needed to account for the perceptual decoding of syllabic spread cues: thus, as shown, transitions and murmurs (and, presumably, releases) are evaluated simultaneously and complementarily in a way that a cross-category maximum to minimum perceptual effect for transitions ([ɲ]>[n]>[ŋ]) corresponds invariably to a minimum to maximum effect for murmurs ([ɲ]<[n]<[ŋ]). Such reciprocal correlation conforms to the existence of a trading relation between transitions and murmurs for [ŋ] and [n] with [a] in Catalan (section 3.3.C.) and a defined compensatory effect between strongly positive [ɲ] transitions and perceptually irrelevant [ɲ] murmur. Further evidence for simultaneous processing of nasal cues at the syllabic level according to vowel quality has been reported to be true for [ŋ] preceded by [æ], [ɛ] (section 2.2.C.).

A perceptual model that allows this sort of auditory analysis of transition ranges to occur presupposes reference to a basic set of articulatory gestures but is not compatible with a feature recognition model based upon the auditory detection of invariant short-term spectral properties. Blumstein and Stevens (1979) report properties of this sort at a 6 msec window release period of [m] (diffuse-falling frequency-amplitude template) vs. [n] (diffuse-rising frequency-amplitude template) in initial position. Frequency-amplitude spectral characteristics at the release can be hardly thought to be "primary place cues" since unreleased nasals are equally common and occurring releases may be perceptually weak, almost indistinguishable; moreover, murmur spectra can hardly provide such cues, given their low amplitude component, high variability and particular pole-zero structure. In fact, transitions have been found in my experiment to be the best place cues in combination with appropriate murmurs: overall, gliding F2-F3 transitions have given more than 95% of [ɲ] judgments and optimal transition-murmur combinations 80% of [n] and [ŋ] responses. While, in order to handle these aspects, the perceptual model proposed by Stevens and his coworkers can be shown to be too simple and limited, it becomes too complex and complicated on other grounds. Thus, contrary to what has been suggested by Stevens (1975), examination of long stretches of acoustic data (presumably hundreds of msec long) before phonetic feature decoding begins is not needed in the case of diphthong-like spectral nuclei such as palatal and palatalized articulations: short pre-closure transitions are satisfactory cues for [ɲ] with [a] and can be processed in the same way as those for [n] and [ŋ].

I have argued for a perceptual process of nasals based upon the simultaneous integration of acoustic cues according to demands imposed by the production and auditory systems. But, given [a] and other possible vocalic environments, what is the articulatory basis for this integration process? Little perceptual data on the identification of nasals of different places of articulation in different vocalic environments is available. The only systematic approach is that of Henderson (Note 1). Henderson's data and evidence provided in this paper support the view that the perceptual interpretation of transitions and murmurs for nasals preceded by different vowels is similar to the integration of burst and transitions for non-nasal stops of the same place of articulation in CV environments (Dorman et al., 1977). Thus [ŋ] murmurs are optimal cues and [ŋ] transitions very weak cues for [i], [a], [ɔ], [o], [u]; for [e], [ɛ] the role of transitions becomes more relevant, while for [e]--but not for [ɛ]--that of the murmur decreases. For [n], transitions are very effective cues with back vowels but very weak with [i] and [e] while murmur is, complementarily, a better cue for [i], in accordance with earlier findings (Hecker, 1962; Nakata, 1959; Ohala, 1975). General effects for [n] are to be expected also for [ɲ] and ought to be correlated with long transition duration with back vowels vs. short transition duration and little excursion range with high vowels (see, for Polish, Jassem, 1964).

While the perceptual relevance of vowel-to-consonant transition ranges for nasals accords well with data from Dorman et al. (1977), that of murmurs is only consistent for [ŋ] with all vowels but [ɛ] and for [n] with [i]. Differently from alveolar bursts, for [n], strong murmur effect is found in the case of [o], [ɛ], [u] and less for [a], [e], [ɔ]. Such findings about the category identification role of murmurs in different vocalic environments suggest that the interpretation strategy used by listeners in associating murmur and articulatory event differs from that proposed by Dorman et al. for burst and front cavity size. In the case of final nasals, different articulatory conditions argue for different integration strategies: no burst is present and release is weak and unnecessary; spectral continuity cannot be expected between oral transition endpoints and oro-nasal murmur concentrations characterized, moreover, by a low perceptual relevance of the mid spectral regions; finally, energy concentration and crucial place information in nasal murmurs are dependent on the size characteristics of the oro-nasal system behind the tongue constriction point (back cavity). A plausible integration model for Henderson's data would associate the back cavity for the nasal consonant with the overall front-back cavity system for the vowel so that the perceptual effectiveness of the murmur would depend on the degree of similarity between back nasal cavity and front or back cavity size appropriate for the vowel. Such a model would predict perceptual relevance for [ŋ] murmur with the back cavity size of coarticulated back and front vowels, for [m] murmurs with the considerable front cavity size of back vowels, and for [ɲ] murmur with the wide pharyngeal pass of palatal vowels. For [n] murmur there would be integration with the overall tract system of a mid vowel and the front cavity of [i]. Obviously, more experimental evidence is needed. In any case, data from Henderson (Note 1) and experiments reported in this paper show that transitions and murmurs, analogously to bursts and transitions, are equivalent and complementary.

REFERENCE NOTE

1. Henderson, J. *On the perception of nasal consonants*. Unpublished Generals Examination paper, University of Connecticut, 1978.

REFERENCES

Ali, A., Gallagher, T., Goldstein, J., & Daniloff, R. Perception of coarticulated nasality. *Journal of the Acoustical Society of America*, 1971, *49*, 538-540.

Bjuggren, G., & Fant, G. The nasal cavity structures. *Royal Institute of Technology, STL-QPSR*, 1964, *4*, 5-7.

Blumstein, S. E., & Stevens, K. N. Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. *Journal of the Acoustical Society of America*, 1979, *66*, 1001-1017.

Carlson, R., Granström, B., & Pauli, S. Perceptive evaluation of segmental cues. *Royal Institute of Technology, STL-QPSR*, 1972, *1*, 18-24.

Cooper, F. S., Delattre, P. C., Liberman, A. M., Borst, J. M., & Gerstman, L. J. Some experiments on the perception of synthetic speech sounds. *Journal of the Acoustical Society of America*, 1952, *24*, 597-606.

Dalby, J., & Port, R. Radial trajectories in F2XF3 plane as place invariants. *Research in Phonetics*, 1980, *1*, 201-216.

Delattre, P. The physiological interpretation of sound spectrograms. *Publications of the Modern Language Association of America*, 1951, *66*, 864-875.

Delattre, P. Les attributs acoustiques de la nasalité vocalique et consonantique. *Studia Linguistica*, 1954, *8*, 103-109.

Delattre, P. Les indices acoustiques de la parole. *Phonetica*, 1958, *2*, 226-251.

Delattre, P. Divergences entre nasalités vocalique et consonantique en français. *Word*, 1968, *24*, 64-72.

De Mori, R., Gubrynowicz, R., & Laface, P. Inference of a knowledge source for the recognition of nasals in continuous speech. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 1979, *ASSP-27*, *5*, 538-549.

Derkach, M., Fant, G., & Serpa-Leitão, A. de Phoneme coarticulation in Russian hard and soft VCV-utterances with voiceless fricatives. *Royal Institute of Technology, STL-QPSR*, 1970, *2-3*, 1-7.

Dorman, M. F., Studdert-Kennedy M., & Raphael, L. J. Stop-consonant recognition: Release burst and formant transitions as functionally equivalent, context-dependent cues. *Perception & Psychophysics*, 1977, *22*, 109-122.

Dukiewicz, L. *Polskie gloski nosowe* (Analiza akustyczna). Warsaw: Polska Akademia Nauk, 1967.

Eek, A. Acoustical description of the Estonian sonorant types. *Estonian Papers in Phonetics*, 1972, 9-35.

Fant, G. *Acoustic theory of speech production*. The Hague: Mouton, 1960.

Fant, G. Descriptive analysis of acoustic aspects of speech. *Logos*, 1962, *5*, 3-17.

Fant, G. Auditory patterns of speech. In W. Wathen-Dunn (Ed.), *Models for the perception of speech and visual form*. Cambridge, Mass.: MIT Press, 1967, 111-125.

Fant, G. Acoustic description and classification of phonetic units. *Speech sounds and features*. Cambridge, Mass.: MIT Press, 1973, 32-83.

Fant, G. Perspectives in speech research. *Royal Institute of Technology, STL-QPSR*, 1980, *2-3*, 1-16.

Ferrero, F., Genre, A., Boë, L. J. & Contini, M. *Nozioni di Fonetica Acustica*. Torino: Edizioni Omega, 1979.

Ferrero, F., Vagges, K., Righini, G., & Pelamatti, G. M. Un sistema di sintesi dell'italiano: primi risultatti. *Rivista Italiana di Acustica*, 1977, *1*, 33-48.

Fischer-Jørgensen, E. Acoustic analysis of stop consonants. *Miscellanea Phonetica*, 1954, *2*, 42-59.

Fujimura, O. Analysis of nasal consonants. *Journal of the Acoustical Society of America*, 1962, *34*, 1865-1875.

Fujimura, O. Formant-antiformant structure of nasal murmurs. *Proceedings of the Speech Communication Seminar* (1962). Vol. I. Stockholm: Royal Institute of Technology, Speech Transmission Laboratory, 1963, 1-9.

Fujimura, O. Remarks on stop consonants. Synthesis experiments and acoustic cues. In L. L. Hammerich, R. Jakobson, & E. Zwirner (Eds.), *Form and substance*. Denmark: Akademisk Forlag, 1971, 221-232.

Fujimura, O., & Lindqvist, J. Sweep-tone measurements of vocal-tract characteristics. *Journal of the Acoustical Society of America*, 1970, *49*, 541-557.

Garcia, E. The identification and discrimination of synthetic nasals. *Haskins Laboratories Status Report on Speech Research*, 1966, *SR-7/8*, 3.1-3.16.

Garcia, E. Labelling of synthetic nasals. *Haskins Laboratories Status Report on Speech Research*, 1967, *SR-9*, 4.1-4.17. (a)

Garcia, E. Discrimination of three-formant nasal-vowel syllables. *Haskins Laboratories Status Report on Speech Research*, 1967, *SR-12*, 143-153. (b)

Green, P. S. Consonant-vowel transitions. A spectrographic study. *Studia Linguistica*, 1959, *13*, 57-105.

Halle, M., Hughes, G. W., & Radley, J.-P. A. Acoustic properties of stop consonants. *Journal of the Acoustical Society of America*, 1957, *29*, 107-116.

Harris, K. S., Hoffman, H. S., Liberman, A. M., Delattre, P. C., & Cooper, F. S. Effect of third-formant transitions on the perception of the voiced stop consonants. *Journal of the Acoustical Society of America*, 1958, *50*, 122-126.

Haskins Laboratories Quarterly Progress Report. Research Study on Reinforcement of Speech. Number: Eight (1953). Haskins Laboratories.

Haskins Laboratories Quarterly Progress Report. Research Study on Reinforcement of Speech. Number: Eleven (1954). Haskins Laboratories.

Haskins Laboratories Quarterly Progress Report. Research Study on Reinforcement of Speech. Number: Thirteen (1954). Haskins Laboratories.

Hattori, S., Yamamoto, K., & Fujimura, O. Nasalization of vowels in relation to nasals. *Journal of the Acoustical Society of America*, 1958, *30*, 267-274.

Hecker, M.H. Studies of nasal consonants with an articulatory speech synthesizer. *Journal of the Acoustical Society of America*, 1962, *34*, 179-188.

House, A. S. Analog studies of nasal consonants. *Journal of Speech and Hearing Disorders*, 1957, *22*, 190-204.

Jassem, W. The acoustics of consonants. In A. Sovijärvi & P. Aalto (Eds.), *Proceedings of the Fourth International Congress of Phonetic Sciences*. The Hague: Mouton, 1962, pp. 50-72.

Jassem, W. A spectrographic study of Polish speech sounds. In D. Abercrombie, D. B. Fry, P. A. D. MacCarthy, N. L. Scott, & J. L. M. Trim (Eds.), *In honour of Daniel Jones*. London: Longmans, Green, 1964, 334-348.

Jassem, W. *Podstawy fonetyki akustycznej*. Warsaw: Polska Akademia Nauk, 1973.

Kacpranski, R. P. Spectral analysis of German nasal consonants. *Phonetica*, 1965, *12*, 165-170.

Kacprowski, J., & Mikiel, W. Simplified rules for parametric synthesis of nasal and stop consonants in C-V syllables by means of the "terminal-analog" speech synthesizer. *Acoustica*, 1965, *16*, 356-364.

Klatt, D. H., & Shattuck, S. R. Perception of brief stimuli that resemble rapid formant transitions. In G. Fant & M. A. A. Tatham (Eds.), *Auditory analysis and perception of speech*. New York: Academic Press, 1975, 293-301.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. Perception of the speech code. *Psychological Review*, 1967, *74*, 431-461.

Liberman, A. M., Delattre, P. C., Cooper, F. S., & Gerstman, L. J. The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs*, 1954, *68*, No. 8.

Magdics, K. *Studies in the acoustic characteristics of Hungarian speech sounds*. Indiana University Publications, Uralic and Altaic Series, 1969, *97*.

Malécot, A. Acoustic cues for nasal consonants: An experimental study involving a tape-splicing technique. *Language*, 1956, *32*, 274-284.

Manrique, A. M. B. de, Gurlekian, J. A., & Massone, M. I. Función de las propiedades acústicas en el reconocimiento de las consonantes nasales y líquidas españolas. *Informe XIII del Laboratorio de Investigaciones Sensoriales*. Buenos Aires, 1980, *13*, 6.

Mártony, J. The role of formant amplitudes in synthesis of nasal consonants. *Royal Institute of Technology, STL-QPSR*, 1964, *3*, 28-31.

Mascaró, J. *Catalan phonology and the phonological cycle*. Bloomington: Indiana University Linguistics Club, 1978.

Massone, M. I. Estudio acústico de las consonantes españolas nasales y líquidas. *Informe XIII del Laboratorio de Investigaciones Sensoriales*. Buenos Aires, 1980, *13*, 5.

Mattingly, I. *Synthesis by rule of General American English*. Supplement to Haskins Laboratories Status Report on Speech Research, 1968.

Mattingly, I., Pollock, S., Levas, A., Scully, W., & Levitt, A. Software synthesizer for phonetic research. *Journal of the Acoustical Society of America*, 1981, *69*, S83. (Abstract)

Miller, J. L., & Eimas, P. D. Studies on the perception of place and manner of articulation: A comparison of the labial-alveolar and nasal-stop distinctions. *Journal of the Acoustical Society of America*, 1977, *61*, 835-845.

Nakata, K. Synthesis and perception of nasal consonants. *Journal of the Acoustical Society of America*, 1959, *31*, 661-666.

Nord, L. Perceptual experiments with nasals. *Royal Institute of Technology, STL-QPSR*, 1976, *2-3*, 5-8.

Ohala, J. J. Phonetic explanations for nasal sound patterns. In Ch. A. Ferguson, L. M. Hyman, & J. J. Ohala (Eds.), *Nasalfest*, 1975, 289-316.

Pickett, J. M. Some acoustic cues for synthesis of the /n-d/ distinction. *Journal of the Acoustical Society of America*, 1965, *35*, 474-477.

Potter, R. K., Kopp, G. A., & Green, H. C. *Visible speech*. New York: D. Van Nostrand Inc., 1947.

Romportl, M. Zur akustischen analyse und Klassifizierung der nasale. *Studies in Phonetics*. The Hague: Mouton, 1973, 78-83.

Stevens, K. N. Feature detection and auditory segmentation: Consonant perception. In G. Fant & M. A. A. Tatham (Eds.), *Auditory analysis and perception of speech*. New York: Academic Press, 1975, 191-195.

Studdert-Kennedy, M. Universals in phonetic structure and their role in linguistic communication. In T. H. Bullock (Ed.), *Recognition of complex acoustic signals*. Berlin: Dahlem Konferenzen, 1977, 37-48.

Takeuchi, S., Kasuya, H., & Kido, K. A method for extraction of the spectral cues of nasal consonants. *Journal of the Acoustical Society of Japan*, 1975, *31*, 739-740.

Vagges, K., Ferrero, F. E., Caldognetto-Magno, E., & Lavagnoli, C. Some acoustic characteristics of Italian consonants. *Journal of Italian Linguistics*, 1978, *3*, 69-85.

Wang, W. S.-Y., & Fillmore, Ch. J. Intrinsic cues and consonant perception. *Journal of Speech and Hearing Research*, 1961, *4*, 130-136.

Weinstein, C. J., McCandless, S. S., Mondshein, L. F., & Zue, V. W. A system for acoustic-phonetic analysis of continuous speech. *IEEE Transactions on Acoustics, Speech and Signal Processing* , 1975, *ASSP-23*, *1*, 54-67.

Zee, E. Effect of vowel quality on perception of post-vocalic nasal consonants in noise. *Journal of Phonetics*, 1981, *9*, 35-48.

FOOTNOTES

[1] I represent with [ɲ] palatal as well as palatalized nasal stops. Velar nasal stops are consistently represented with [ŋ] independently of their phonological status.

[2] Research is to be done on the relative differences in pharynx cavity size and size of velum opening among nasals of different place of articulation. It deserves to be seen to what extent the acoustic structure and the perceptual role of their murmur formants are affected by those differences.

[3] The reliability of the results remains open to the objections raised in section 3.3.C. Furthermore, possible bias effects can be related to the fact that all experiments were forced-choice. Also, the fact that American English subjects in Malécot's experiment gave 80% of [ŋ] responses for original [ŋ] while 100% of [m] and [n] responses for original [m] and [n], respectively, suggests response bias effects against the velar correlate. That American English speakers identify [ŋ] after [a] very reliably has been shown with

original natural speech stimuli by Henderson (Note 1) (100%) and Zee (1981) (96%).

[4] I call transition range value the amount of frequency contrast between steady state vowel and starting point or endpoint value of an adjacent formant transition. It is expressed in Hz and can be positive (+), negative (-), or null (=).

[5] Acoustic measurements of zeros were inferred from frequency areas that show a major reduction in the magnitude of the energy envelope as observed in spectrographic spectral sections. A final decision about zero frequency values to be included in the synthetic speech patterns used for perceptual testing was also reached on the basis of measurements reported in the literature as well as well-accepted observations on formant-antiformant spectral characteristics of nasal murmurs corresponding to different place categories (see Sections 2.1.B. and 2.2.A., and Table 1).

[6] Bandwidth values were estimated by measuring the distance between two points at the right and left side of the spectrum envelope, equally located 3dB below the peak level.

[7] The mismatch between perceptual stimuli and frequency values corresponding to the [ɲ] production space (Figure 7) did not affect the quality of [ɲ] judgments, thus suggesting that subjects perceive a positive F2 transition as a palatal cue when pointing to a critical high locus.

SPEECH PRODUCTION CHARACTERISTICS OF THE HEARING IMPAIRED*

Mary Joe Osberger+ and Nancy S. McGarr++

I. INTRODUCTION

One of the most devastating effects of congenital hearing loss is that normal development of speech is often disrupted. As a consequence, most hearing-impaired children must be taught the speech skills that normal-hearing children readily acquire during the first few years of life. Although some hearing-impaired children develop intelligible speech, many do not. For many years, it was believed that profoundly hearing-impaired children were incapable of learning to talk. Carrying this belief to the extreme, Froeschels (1932) even suggested that all deaf children exhibited some behavior problems, "due to the fact that the profuse motor release connected with speech is impossible in their case" (p. 97).

Within the last decade, advances have been made in studying the speech of the hearing impaired. This is largely due to the development of sophisticated processing and analysis techniques in speech science, electrical engineering, and computer science that have increased our knowledge of normal speech production. In turn, these technological advances have been applied to the analysis of the speech of the hearing impaired, and also to the development of clinical assessment and training procedures.

The oral communication skills of hearing-impaired children have long been of concern to educators of the hearing impaired, speech pathologists, and audiologists because the adequacy of such skills can influence the social, educational, and career opportunities available to these individuals. Since the introduction of PL 94-142 and the emphasis on mainstreaming, there is an even greater likelihood that many professionals will need to learn about, or upgrade their knowledge of, the speech of hearing-impaired children. The intent of this chapter is to provide the clinician, student, and researcher with a comprehensive description of the speech characteristics of this

---

*To appear in N. Lass (Ed.), Speech and language: Advances in basic research and practice (Vol. 8). New York: Academic Press.

+Human Communication Research Laboratories, The Boys Town Institute for Communication Disorders in Children, Omaha, Nebraska.

++Also Communication Sciences Laboratory, City University of New York.

Acknowledgement. Preparation of this manuscript was supported in part by NINCDS Grant NS-16247 and by NIH Biomedical Research Support Grant #1S07RR05834 to Boys Town Institute for Communication Disorders in Children; and NINCDS Grants NS-13617 and NS-13870, and NIH Biomedical Research Support Grant RR-05596 to Haskins Laboratories. The authors wish to thank Drs. Katherine Harris, Ray Kent and Charles Watson for their helpful comments.

population. It is assumed that the reader has some familiarity with the effects of congenital hearing loss on speech and language development, and has some exposure to acoustic and articulatory phonetics. It should be noted that most of the information available about the hearing impaired is concerned with children with severe and profound sensorineural hearing losses (losses of 70 dB HTL or greater). In comparison, relatively little is known about the speech of hard-of-hearing children (losses less than 70 dB HTL). It is for this reason that most of the chapter is devoted to children who are severely and profoundly hearing impaired.

In order to present an in-depth coverage of speech production processes, we have opted to discuss language skills only in those instances where there is no clear-cut separation between language and speech. Likewise, the auditory skills of the hearing impaired will be discussed only to the extent that factors such as hearing level and auditory capabilities affect speech production skills. The emphasis on speech production is not meant to suggest that an aural/oral teaching method is the only appropriate educational plan for hearing-impaired children. The issues involving educational methodologies are not of primary concern here. Rather, it is the belief of the authors that every hearing-impaired child is entitled to speech training services, even if a realistic goal of such training may be only the development of functional (survival) speech skills. Before optimal teaching strategies can be selected, however, teachers and clinicians must have a thorough understanding of the nature of the problems they are trying to remediate.

II. DEVELOPMENTAL ASPECTS OF THE SPEECH OF THE HEARING IMPAIRED

A. Vocalization Patterns

For many years, it was believed that the vocalization development of hearing and hearing-impaired infants was the same, at least through the babbling stage. After this period, hearing-impaired infants were reported to stop babbling. This notion was based primarily on Mavilya's (1968) data, which showed a marked decrease in the number of vocalizations produced by three congenitally hearing-impaired infants (12-16 weeks old at the start of the study) over a three-month period. Recent data obtained by Stark (in press) do not support the findings of Mavilya. For a group of hearing-impaired infants 15-24 months old, Stark observed an overall increase in rate of vocal output with age. The mean number of vocalizations was also observed to increase as progressively higher levels of vocal output were attained by the infants. In general, the stages of vocalization behavior of the 15- to 24-month-old hearing-impaired infants were similar to those of a group of normal-hearing infants 9-48 weeks of age. An important point that should be made is that the speech behavior of the infants in both the Stark and Mavilya studies was recorded before the children had been fitted with hearing aids. Stark found that the level of vocal development reached by the children before they were fitted with amplification did not appear to predict their later progress in learning speech. The vocal development of some children progressed rapidly after they were given hearing aids, while the vocal development of others did not.

Although Stark found no difference in rate of vocal output between the normal-hearing and hearing-impaired infants, differences in the phonemic repertoire were found between normal infants and hearing-impaired infants who were judged to be at the same level of vocal development. Syllable shape (e.g., CV, VC, CVC, etc.) was similar among the children, but the inventory of vowel- or consonant-like sounds was more limited in the samples and tended to be more stereotyped than those of hearing infants of the same age. Mavilya also observed that the phonemic aspects of the vocalizations of the hearing-impaired infants in her study were different from those reported for infants with normal hearing. Specifically, Mavilya observed a severe delay in the development of consonant sounds in the vocalizations of the hearing-impaired infants, with vowels produced more often than consonants.

In an earlier study, Stark (1967) analyzed the phonemic aspects of the vocalizations of six congenitally hearing-impaired children between the ages of 16 and 19 months before they were fitted with hearing aids. Analysis of the infants' vocalizations revealed that the following sounds were used by all six babies: (1) a low front vowel, such as /æ/; (2) a neutral mid-vowel or schwa; (3) an aspirant /h/, which could precede or follow vowel sounds; (4) a syllabic nasal consonant usually identified as /m/, and (5) a glottal stop. An interesting observation made during this study was that the emotive vocalizations of the hearing-impaired infants, such as whimpering, sighing, crying, and laughing, did not sound deviant, and therefore this aspect of vocal behavior did not provide diagnostic information about the hearing status of infants.

In summary, the results of Stark's (1967, in press) research do not support the belief that hearing-impaired infants simply cease vocalizing after the babbling stage. Differences between the vocalizations of normal-hearing and hearing-impaired infants do emerge at an early age, but the differences are seen in phonemic production rather than quantity of vocal output as suggested by Mavilya (1968).

B. Speech Sound Inventories

Phonetic inventories have been obtained from the spontaneous samples of hearing-impaired children ranging from eleven months to seven years of age (Carr, 1953; Lach, Ling, Ling, & Ship, 1970; Stark, in press; Sykes, 1940; West & Weber, 1973). Although these studies report differences in the frequency of specific vowel sounds in the samples of hearing-impaired children studied, the pattern of vowel production is remarkably similar. The vowels most commonly used by young hearing-impaired children include the central vowels /ʌ, ə/ and the low front vowels /ɛ, æ/. The extreme high vowels /i, u/ occurred relatively infrequently in the children's samples. The exception to this pattern was reported by Carr (1953), whose five-year-old hearing-impaired subjects used a wider range of vowels than noted above. There is some evidence that this pattern of vowel usage changes over time. For example, Lach et al. (1970) found that over a one-year period, young hearing-impaired children, 11-32 months of age, who were enrolled in a preschool program, tended to shift from the frequent use of the schwa vowel to other vowels, with the greatest increase in usage observed for /I/. Carr (1953) also compared the relative frequency of each vowel type in the speech of five-year-old hearing-impaired children to that of hearing children and noted that the hearing-impaired children used vowels in a manner and degree similar to hearing infants of 11 to 12 months of age. The hearing-impaired children were also found to use vowel sounds more often than consonant sounds. In another study, Sykes (1940) found that 4- to 7-year-old hearing-impaired children produced almost half of their vowel sounds in isolation and not in combination with a consonant.

Analyses of consonant production have shown that young hearing-impaired children produce front consonants /b, p, m, w/ more often than they produce back consonants (Carr, 1953; Lach et al., 1970; Sykes, 1940), and they have been found to use front consonants with greater frequency than do hearing children (Carr, 1953). In a longitudinal study, Lach et al. (1970) analyzed consonant usage by manner of production. Before the children began a preschool program, 66% of all consonants produced were glottal sounds, and approximately 25% of the sounds were nasal consonants. After one year in the program, the glottal sounds were used only 44% of the time. There was also a large increase in the usage of plosives and semi-vowels, due primarily to an increased use of /b/ and /w/. Fricatives and affricates were used only rarely, even after one year of training. With only one exception, all children produced a significantly greater number of consonants and vowels after one year of training, with a concomitant increase in the consonant-to-vowel ratio.

C. Phonemic and Phonologic Skills

A limitation of the simple sound-type inventories, which were discussed above, is that no information is provided on the phonological usage of the speech segments. A tally of consonants and vowels does not reveal whether or not the phonemes were used appropriately. To overcome this limitation, investigators have begun to perform phonemic, phonological, and linguistic analyses of hearing-impaired children's speech (Oller, Jensen, & Lafayette, 1978; Oller & Kelly, 1974; Stoel-Gammon, in press; West & Weber, 1973).

Recently, a comprehensive study was performed by Stoel-Gammon (in press) in which cross-sectional and longitudinal data were obtained on phonological acquisition by hearing children, 1.5 to 3.10 years of age, and hearing-impaired children, 2.4 to 7.3 years of age. The cross-sectional data showed that, in large part, the patterns of development were similar for the two groups of children, although the rate of development was considerably slower for the hearing-impaired children than for the hearing children. Similar patterns of correct production and error types were present for both groups of children. The set of substitution patterns common to both groups included voicing of initial stops, devoicing of final stops, fricatives, and affricates, and substitution of homorganic stops for fricatives. When errors were common to both groups, they were more frequent in the speech of the hearing impaired than in the speech of the normal-hearing children.

Some differences in the pattern of development between the normal-hearing and hearing-impaired children were also observed in the above study. Errors found to be present only in the hearing-impaired children's speech were: substitution of a glottal stop for the target phoneme, substitution of the palatal fricative /ʃ/ for the affricates /tʃ and dʒ/, and substitution of consonants /h, k, g/ for other non-labial consonants. The only substitution that Stoel-Gammon found to occur in the normal children's productions that did not occur in those of the hearing-impaired was depalatization of /ʃ, tʃ, dʒ/, resulting in a substitution of /s/ for /ʃ/ or /ts/ for /tʃ, dʒ/. The data also showed that the substitutions of the hearing-impaired children deviated farther from the target phoneme with respect to manner and place of production than did the substitutions of the normal children. In addition, the errors of the hearing-impaired subjects also tended to show a larger range of substitution types, for example, /k, g/ for /tʃ/, than those made by the hearing children.

The longitudinal data obtained by Stoel-Gammon revealed that the hearing-impaired children progressed toward correct production of target phonemes at a much slower rate than the normal-hearing children and that there was a much greater range and variation of response types, both within and across subjects. The preliminary data suggested that the hearing-impaired children passed through three developmental stages. In the first stage, the child produced a wide variety of substitutions for the target phoneme. In the second stage, there was a narrowing of the range of substitutions, followed by substitutions with a single sound. In the third stage, the phoneme was produced correctly. Of course, not all hearing-impaired children progress through the third stage, as evidenced by numerous phonetic errors that remain in the speech of many hearing-impaired persons even throughout their adult life.

Additional research is needed in order to delineate the stages of speech acquisition in hearing-impaired children. This information is essential to help us better understand why some children develop intelligible speech and others do not. Although there are data suggesting that hearing-impaired children are simply delayed in phonemic acquisition (Oller et al., 1978; Oller & Kelly, 1974; Stoel-Gammon, in press), we also know that there are differences in the phonology used by hearing children and hearing-impaired children. In fact, there are noticeable differences between the production patterns of the two groups of children at a very early age, and the speech of some hearing-impaired children never progresses beyond the very early stages of development. As we shall see in the following section, the speech production patterns of older hearing-impaired children show many similarities to the patterns of the younger hearing-impaired children. It will also become evident that although in many cases, hearing-impaired children fail to follow rules typical of normal speech, the deviations in their speech show systematic patterns, indicating that they are using a set of phonological rules, even though these rules may differ from those used by normal speakers.

III. ARTICULATORY PATTERNS IN THE SPEECH OF SEVERELY AND PROFOUNDLY HEARING-IMPAIRED CHILDREN

A. Production of Consonants

1. Overview

Perhaps of all the speech production errors characteristic of the severely and profoundly hearing impaired, the area that has received the greatest attention is that involving the articulation of consonants, vowels, and diphthongs. Numerous independent investigations (Hudgins & Numbers, 1942; Markides, 1970; Smith, 1975; McGarr, 1980) have been remarkably consistent in identifying typical articulatory errors in the speech of hearing-impaired children who were trained in many different programs. Most of these investigations are of a descriptive nature; that is, either listener judgments or phonetic transcriptions were used to obtain measurements of intelligibility or to describe the articulatory characteristics of the speech. However, some investigators (Calvert, 1961; Monsen, 1974, 1976b, 1976c; Rothman, 1976) have begun to detail some of the acoustic characteristics of the speech of the hearing impaired (e.g., voice onset time, closure duration, formant frequencies). Acoustic analysis of hearing-impaired speech permits a finer-grained consideration of some aspects of both correct and incorrect productions than would be possible using methods applied in the descriptive literature.

For purposes of organization we will consider the production of suprasegmentals as well as other factors that affect the intelligibility of speech later in this chapter. This section will present information only on the segmental aspects of hearing-impaired children's speech. We will first consider the error patterns detailed in the descriptive literature and then discuss the relevant acoustic data for production of consonants and vowels by hearing-impaired speakers.

2. Consonant Errors

Any comprehensive analysis of the articulatory skills of hearing-impaired children must begin with the classic work of Hudgins and Numbers (1942). These authors studied 192 subjects between the ages of 8-20 years whose hearing losses ranged from moderate to profound. The students read simple sentences. From recordings, teachers of the deaf later evaluated the students' productions for proficiency in articulation as well as rate and rhythm. Error categories were established for consonants, vowels, and diphthongs, and an attempt was made to relate these patterns to speech intelligibility.

Briefly, the articulatory errors were divided into substitutions, omissions, and severe distortions of the intended phoneme as well as the addition of adventitious phonemes or syllables. Among the more common error types involving consonants were confusion of the voiced and voiceless sounds, substitution of one consonant for another, added nasality, misarticulation of consonant blends, misarticulation of abutting consonants, and omission of word initial or word final consonants. This overall pattern of consonant errors has been replicated in numerous studies (Brannon, 1966; Geffner, 1980; Gold, 1978; Levitt, Smith, & Stromberg, 1976; Markides, 1970; Nober, 1967; Smith, 1975), although the actual percentage of errors in any category may vary somewhat from study to study.

a. Voicing errors. Errors of voicing were one of the most frequent types of consonant errors found by Hudgins and Numbers (1942). In subsequent studies, the direction of this error has sometimes been reported as occurring to the voiced member of the pair (Carr, 1953; Heider, Heider, & Sykes, 1941; Millin, 1971; Smith, 1975), and at other times, to the voiceless cognate (Mangan, 1961; Markides, 1970; Nober, 1967).

Smith's (1975) study of 40 severe to profoundly hearing-impaired children enrolled in an oral school for the deaf has been among the most comprehensive since Hudgins and Numbers. The 40 children read sentences containing key words that incorporated the most frequent English phonemes with transition to and from the vowels /i/, /æ/, and /u/ for all places of articulation. Voicing errors were common for these children and most often involved substitutions of the voiced for voiceless member of the pair. Studies by Heider et al. (1941) and Carr (1953) have also reported a tendency for hearing-impaired children to use more voiced than voiceless sounds in their spontaneous speech. Indeed, Millin (1971) suggested that one manifestation of the voiced for voiceless problem is inappropriate phonation evidenced at the beginning or end of an utterance.

This error pattern, voiced for voiceless substitution, is opposite to that found by Markides (1970) who studied 110 British hard-of-hearing and deaf children. The children produced words as part of an articulation test. A common error was substitution of the voiceless cognate for the voiced. Using the Templin-Darley Test of Articulation, Nober (1967) analyzed production of phonemes by 46 severe and profoundly hearing-impaired children. He reported that voiceless phonemes were produced correctly more often than voiced phonemes. Data obtained by Mangan (1961) can also be interpreted to show the difficulty hearing-impaired children have with voicing contrasts. Subjects in this study were reported to devoice final voiced consonants.

Taken together, these studies suggest that coordination of the articulators necessary for voicing contrast is an exceedingly difficult task for hearing-impaired speakers. Recently, some investigators (McGarr & Löfqvist, in press; Whitehead & Barefoot, 1980, among others) have begun to examine the physiological manifestations of some typical errors in the speech of the hearing impaired. Their data suggest that the nature of the voicing error may be far more complex than is suggested by the descriptive literature. In fact, some hearing-impaired speakers fail to coordinate the timing of respiration, phonation, and articulation in attempting to produce voicing contrasts. More will be said about these findings in a later section.

b. *Substitution errors*: *Place of articulation*. Another common articulatory error in the speech of the hearing impaired involves the substitution of one phoneme for another. Frequently, the substitution is to a phoneme with a similar place of articulation. There is general agreement that phonemes produced in the front of the mouth are more often produced correctly than are phonemes produced in the back of the mouth. This makes sense when one considers that the relative visibility of articulatory gestures should be important to hearing-impaired persons for whom there is reduced auditory information.

Substitution errors involving the same place of articulation have been noted in several studies. Nober (1967) analyzed correctly articulated consonants according to place of articulation and then ranked them from highest to lowest scores as follows: bilabials, 59%; labiodentals, 48%; glottals, 34%; linguadentals, 32%; lingua alveolars, 23%; linguapalatals, 18%; and lingua velars, 12%. Similar patterns of correct production have been reported by Smith (1975) and Gold (1978) except these investigators found that sounds produced in the middle of the mouth were more prone to error than were sounds produced in the back of the mouth.

This general trend--better production for more visible phonemes--has been found not only for production of isolated words and sentences (Huntington, Harris, & Sholes, 1968; Geffner & Freeman, 1980; Levitt et al., 1976; Levitt, Stromberg, Smith, & Gold, 1980; Smith, 1975), but also for spontaneous speech (Carr, 1953; Geffner, 1980; Heider et al., 1941).

Some caution should be exercised, however, in interpreting the importance of visibility in and of itself as a key factor in production. Some articulators, such as the lips, although quite visible, are also relatively more constrained and thus permit fewer possibilities for errors than other articulators such as the tongue. Later we shall discuss some physiological data obtained by Huntington et al. (1968) and McGarr and Harris (1980), which is pertinent to this issue.

c. *Substitution errors*: *Manner of articulation*. A common observation arises from an analysis of consonant errors according to place of articulation. Hearing-impaired speakers tend to position their articulators fairly accurately, especially for those places of articulation that are highly visible, but fail to coordinate properly the movement of the articulators (Huntington et al., 1968; Levitt et al., 1976). The type of consonant substitution that occurs in these cases is often described as one resulting from incorrect timing. These errors are also described as involving an inappropriate manner of articulation.

One example of a common error is the nasal-oral substitution. According to Hudgins and Numbers (1942), errors in nasality may be considered to be a segmental problem and also a problem affecting voice quality, although here we are interested primarily in the former. Non-nasal phonemes were reported by Hudgins and Numbers to be nasalized, and nasal consonants were often produced as stops. Similar findings have also been noted by Markides (1970), Smith (1975), and Stevens, Nickerson, Boothroyd, and Rollins (1976).

Other errors in manner of articulation have also been noted. Smith's hearing-impaired children made most errors producing the following: palatal plosives, fricatives, affricates, and the nasal /ŋ/. Glottals were frequently substituted for stops. Fricatives showed a high rate of substitution to, but not from, the plosives. Affricates were never substituted for other consonants, but tended to be substituted by one of their components, usually the plosive component. However, bilabial plosives, the glides, and the fricatives /f/ and /v/ were often produced correctly. Nober's (1967) results also followed the general pattern reported by Smith. Glides were most often correct, followed by stops, nasals, and fricatives. Similar findings were obtained by Geffner and Freeman (1980) for 67 six-year-old severe and profoundly hearing-impaired children attending schools for the deaf throughout New York State.

The articulatory movements for both alveolar and velar sounds are visually obscure. One reason why alveolar sounds may be more prone to error than velar sounds is that more sounds are produced in the middle than in the back of the mouth. Because of this, precise positioning of the articulators is necessary in order to differentiate correctly all the sounds with a medial place of articulation. Thus, greater variability in articulatory placement can be tolerated before the velar sounds are misperceived by the listener. In any event, a consistent finding is that hearing-impaired children correctly produce the highly visible phonemes (i.e., those produced in the front of the mouth) more often than those phonemes that are not articulated with a high degree of visibility (i.e., those produced in the middle or back of the mouth).

d. Omission errors. By far the single most frequently reported error in the speech production of the severely and profoundly hearing impaired is the omission of a phoneme (Hudgins & Numbers, 1942; Markides, 1970; Smith, 1975). Omission of consonants may occur in the initial and/or final position of words, also reported as non-function of releasing or arresting consonants, respectively.

Hudgins and Numbers reported that omission of initial consonants was more common than omission of final consonants. The consonants most frequently omitted from the initial position of words included /h, l, r, y, th, s/. Turning to final consonants, the authors describe several error patterns: dropping of the consonant completely, releasing the consonants into the following syllable, or incomplete production whereby the phoneme loses its dynamic properties and becomes merely a passive gesture. Among the final consonants most frequently omitted in the study by Hudgins and Numbers were /l, t, s, z, d, g, k/. These results are in agreement with those reported by Geffner (1980), who analyzed the spontaneous speech samples of young hearing-impaired children.

Others (Nober, 1967; Markides, 1970; Smith, 1975) have also observed the omission of similar consonants from the speech of hearing-impaired children. In contrast to Hudgins and Numbers, however, these studies reported a greater number of consonants omitted from the final position of words than from either the initial or medial positions.

e. Consonant-cluster errors. Not many investigators have reported data for production of consonant blends. This is surprising since Hudgins and Numbers suggested that these errors had an important and deleterious effect on intelligibility. In their study, these errors involved two forms: one or more components of the cluster were dropped, or an adventitious phoneme, usually the /ə/, was added between the elements. This latter error may be particularly detrimental to the timing or rate and rhythm of speech. Brannon (1966) also found that misarticulation of consonant blends was a significant error in the speech of hearing-impaired children. Smith (1975) tested consonant blends /p, t, k/ and /s/ in the speech production of older hearing-impaired children (13-15 years old). Here again, there was frequent omission of one or more elements of the cluster. In fact, a phoneme in the blend environment was more likely to be omitted than the same phoneme occurring in a non-blend environment.

B. Acoustic Characteristics of Consonant Production

We now turn to a discussion of the acoustic patterns of consonant production. While these consonantal features have been much studied in normal and also in synthetic speech (cf. Borden & Harris, 1980; Pickett, 1980, for a review of this work), there have been far fewer studies of the acoustic characteristics of consonants produced by hearing-impaired speakers. This is in part because spectral measurements of hearing-impaired speech are particularly difficult to make, either because of the mismatch between spectrograph filter and fundamental frequency (cf. Huggins, 1980), or because of source function abnormalities.

In normal speech production, the acoustic consequences of consonant production are complex and spread over a period of time. They involve differences in the sound source and the spectral composition of the signal. For example, in the production of a voiceless fricative in a vocalic environment (e.g., VCV, "I see"), the sound source is changed from a periodic to an aperiodic one, and then back to the periodic source. Similarly, a voiceless aspirated stop in a similar VCV environment (e.g., "a pie") is associated with the following sequence of source changes: periodic voicing during the preceding vowel, silence during the consonantal closure, transient noise, aspiration noise, and periodic voicing during the vowel. In addition to being spread across time, the acoustic attributes of many consonants often involve short-term spectral changes, where high frequency components play an important role. Examples of such attributes are release bursts and formant transitions for stop consonants, and spectra and transition for fricatives. These characteristics provide considerable information on the identity of segments. In the speech of the hearing impaired, acoustic analysis of consonant production has been made only for voice-onset-time (VOT), formant transition, or closure and constriction duration, and these patterns give ample evidence of the great perceptual difficulty that listeners to the speech of the hearing-impaired experience.

1. Voiced-Voiceless Distinction

At the acoustic level, contrasts such as "voiced" versus "voiceless" or "aspirated" versus "unaspirated" are manifested as complexes of acoustic cues (Slis & Cohen, 1969). In the classic study of Lisker and Abramson (1964), release of the oral occlusion relative to the onset of glottal pulsing (i.e., voice-onset-time or VOT) was the salient cue that distinguished voiced from voiceless stops. As was previously discussed, errors in voicing are common in the speech of the hearing impaired. Some acoustic studies of their speech provide evidence that a lack of voice-onset-time contributes to the perception of the voiced-voiceless confusion.

Perhaps the most careful study in this area has been conducted by Monsen (1976b). Spectrographic measurements of VOT were made of word-initial stops /p, t, k/ and /b, d, g/, produced by 36 profoundly hearing-impaired children. Some of the children distinguished the cognates in the normal manner. VOT values were longer for the voiceless than voiced segments and VOT contrasts were longer for velars than for alveolars and bilabials, respectively. However, most of the hearing-impaired speakers did not observe the voiced-voiceless distinction and deviated from normal speakers in a similar way. Typically, voice-onset-time values for voiceless segments were lower than those for voiced, and also overlapped with the measurements for voiced. This pattern was noted for /p-b/ and /t-d/, although measurements for /k-g/ were more complex. Furthermore, these subjects did not distinguish VOT among stops based on place of articulation. Hearing-impaired speakers who observed the voiced-voiceless distinction typically had high speech intelligibility, probably because they were capable of producing other aspects of speech normally as well. Hearing-impaired speakers who did not observe these contrasts tended to collapse the voiced-voiceless categories, producing most segments as voiced. These speakers were considerably less intelligible than those who produced the voicing distinction.

Findings similar to Monsen's have been reported in the earlier work of Calvert (1961) and Irvin and Wilson (1973), and more recently as part of measurements made in studying the acoustic and articulatory correlates of the speech of the hearing impaired (Mahshie, 1980; McGarr & Löfqvist, in press; Stein, 1980). In the McGarr and Löfqvist study, the authors noted that VOT values for some of their hearing-impaired speakers fell in the range of 20-30 msec, which is close to the perceptual boundary where shifts in the perception of voicing have been shown to occur. This may be one reason why listeners to the speech of the hearing impaired have difficulty making judgments of particular phonetic segments. We will return to these physiological studies later.

2. Formant Patterns of Transition

Hearing-impaired speakers have often been described as having difficulty in moving their articulators correctly from one phoneme to the next (Calvert, 1961; John & Howarth, 1965; Martony, 1966; Smith, 1975). One manifestation of this problem at the acoustic level is distortion of formant frequency transitions.

Changes in the formant frequencies, particularly the direction, extent, and duration of the second formant transition, have been shown to be important acoustic cues for the place of articulation (Delattre, Liberman, & Cooper, 1955; Liberman, Delattre, Gerstman, & Cooper, 1956). As discussed above, hearing-impaired speakers characteristically produce many errors involving the place of articulation.

While there have been only a few acoustic analyses of formant transition of hearing-impaired speakers, these studies are nonetheless in general agreement (Martony, 1966; Monsen, 1976c; Rothman, 1976). In general, this work showed that formants were exceedingly short in duration or missing altogether, that the extent of the frequency range of the transitions was limited in part because the formant frequencies for vowels were greatly neutralized, and that transitions varied little with respect to phonetic context. Also, the slope of the transitions frequently remained fairly flat when either a rising or falling pattern was dictated. Thus, F2 transitions in the speech of the hearing-impaired may be reduced in both duration and frequency range. These patterns, together with deviations in the steady state formant frequencies for vowels (to be discussed later), suggest that hearing-impaired speakers have reduced articulatory movement and an absence of the coarticuatory effects observed in the speech of normals.

C. Production of Vowels and Diphthongs

1. Overview

Hudgins and Numbers (1942) were again among the first investigators to study the production of vowels and diphthongs systematically in the speech of the hearing impaired. They classified the errors according to five major types:

1. Substitution of one vowel for another
2. Neutralization of vowels
3. Diphthongization of vowels
4. Nasalization of vowels
5. Errors involving diphthongs: either the diphthong was split into two distinctive components, or the final member of the diphthong was dropped.

In this study, substitutions and neutralization of vowels, and difficulty with the production of diphthongs were among the most common errors. Essentially the same pattern has been replicated in other studies of hearing-impaired speakers regardless of whether the vowel was produced in a CVC framework (Angelocci, Kopp, & Holbrook, 1964; Calvert, 1961), in test words (Geffner, 1980; Mangan, 1961; Markides, 1970; Nober, 1967), or in sentences (Smith, 1975).

There is also agreement concerning the frequency of vowel versus consonant errors. Overall, fewer errors in vowel production have been reported, although it should be noted that this finding may be influenced by variables in both speaker production and listener perception. For example, Brannon (1966) claimed that vowels were in fact easier for hearing-impaired speakers to produce than consonants, since vowels were supposed to require less precise articulatory position. Perceptually, Hudgins and Numbers (1942) and later Monsen (1976c) suggested that listeners tolerate a greater degree of distortion in vowels than in consonants, hence the report of fewer vowel errors. Furthermore, acoustic information conveyed in the vocalic position of the stimulus also provides information of consonants, and thus, if erroneous (as we will discuss later), may directly affect the perception of the consonant. In general, it should also be noted that fewer vowels than consonants are produced in running speech, thus there is less opportunity for error.

2. Vowel Errors

Traditional classification schemes for vowels employ such categories as tongue position (high-low, front-back), tongue tension (tense-lax), and degree of lip rounding. These refer to articulatory events and are important to our subsequent discussion of the acoustic characteristics of vowels. In general, hearing-impaired speakers have been found to produce back vowels correctly more often than front vowels (Boone, 1966; Geffner, 1980; Mangan, 1961; Nober, 1967; Smith, 1975) and low vowels correctly more often than those with mid or high tongue positions (Geffner, 1980; Nober, 1967; Smith, 1975). In fact, Boone (1966) suggested that hearing-impaired speakers tend to keep their tongue retracted in a low back position. In contrast, Stein's (1980) cinefluographic study of vowels produced by hearing-impaired speakers showed "fronting" of back vowels.

With respect to errors of substitution, hearing-impaired speakers often fail to make the tense-lax distinction (Mangan, 1961; Monsen, 1974; Smith, 1975), although there is evidence to the contrary (Hudgins & Numbers, 1942; Markides, 1970). The commonly observed error of neutralization, a problem akin to substitution, has been noted in the descriptive literature (Heider et al., 1941; Markides, 1970; Smith, 1975) as well as in acoustic studies (Angelocci et al., 1964; Monsen, 1976a, 1978). This work suggests that the hearing-impaired speaker tends to produce vowels with a pattern appropriate for the neutral vowel /ə/. This error has implications at the segmental as well as the suprasegmental level since, in the latter case, the syllable is shortened and often not given the appropriate stress.

Other commonly reported errors in vowel production include inappropriate nasalization of vowels (Martony, 1966; Stevens et al., 1976) and diphthongization of pure vowels (Boone, 1966; Markides, 1970; Smith, 1975). With the exception of Hudgins and Numbers (1942), very little additional data have been collected on production of diphthongs, the error patterns reported being essentially the same (Levitt et al., 1980; Nober, 1967).

D. Acoustic Characteristics of Vowels

The acoustic characteristics of vowels and diphthong production have been studied in great detail in normal speech production, but again there is little in the speech of the hearing impaired. We will concentrate here primarily on studies of vowel formants, leaving a discussion of timing characteristics (i.e., duration) and segmental influences on fundamental frequency until later in the chapter.

The formant frequencies, especially the first (F_1) and second (F_2) formants, are traditionally used to provide an acoustic description of vowels. Usually, these formant values are plotted against each other, and the data points for each vowel cluster into fairly distinctive regions (cf. Peterson & Barney, 1952). Interestingly, the acoustic vowel plot of F_1 and F_2 closely resemble the articulatory vowel map. Although the relationship between acoustic and articulatory correspondence is not simple, it has been suggested that F_1 (which increases and then decreases as the vowels go from /i/ to /u/) represents tongue height, and that F_2 (which decreases from /i/ to /u/) represents the constriction of the tongue in the front-back plane. Of course, events such as degree of lip rounding, pharyngeal constriction, as well as individual speaker differences must also be considered.

Analysis of spectrograms of this population is not without problems. In many cases there are extra harmonics and the fundamental frequency (Fo) of hearing-impaired speakers may often be quite high. This may create a mismatch between the source and the bandwidth of the spectrogram filter used in analyzing the acoustic signal. This problem is similar to that faced in the spectrographic analysis of young hearing children's speech (cf. Huggins, 1980). Spectrographic analysis of the speech of the hearing impaired is further complicated by pertubations in the source, inappropriate management of intensity, and/or inappropriate nasalization that introduces zeros into the frequency domain. This often precludes easy and straightforward analysis. Some of these problems may be circumvented by the use of digital analysis techniques such as Linear Predictive Coding (LPC). Even with the use of LPC, determination of formant frequency location may still be difficult.

There have been several studies examining the acoustic characteristics of vowels produced by hearing-impaired children using spectrographic analysis (Angelocci et al., 1964; Monsen, 1974, 1978), and one study in which the speech was digitized and subjected to LPC analysis (Osberger, Levitt, & Slosberg, 1979). Besides instrumentation differences, these studies also differ in that the latter work includes only productions perceived as correct in hearing-impaired children's speech, while the other studies are not clear with respect to this point. Nonetheless, the results of these studies show that the formant frequencies of deaf children's vowels tend toward that of the neutral vowel /ə/. This result is of further interest since the hearing-impaired subjects in both the Monsen and the Osberger et al. studies produced vowels in sentence context, while subjects in Angelocci et al. produced vowels in CVC monosyllables. The data from these studies are interpreted to suggest that hearing-impaired speakers use a restricted amount of tongue movement to achieve vowel differentiation. Indeed, several investigators (Angelocci et al., 1964; Martony, 1968) have suggested that differences in vowels produced by hearing-impaired speakers are achieved primarily by means of variation in fundamental frequency.

In addition to reduced phonological space for all vowels and extensive overlapping of vowel areas, Monsen (1976a) also noted that the second formant of vowels remained around 1800 Hz rather than varying as different vowels were articulated. This "immobility of F_2" not only deleteriously affects perception of the vowel but also interferes with transmission of consonantal information. The difficulty with F_2 is not surprising, since many hearing-impaired speakers have residual hearing only in the frequency range of F_1 and not in the range of F_2. Another factor is that the front-back tongue movement associated with second formants is impossible for a deaf student to observe. Articulatory movements such as jaw movement associated with F_1 change are certainly more visible.

Very little is known about the acoustic aspects of diphthong production in the hearing impaired. Monsen (1976d), using spectrographic measurements of the diphthong /aI/, has classified deviant acoustic patterns on the basis of frequency change during production of the diphthong. One deviant pattern is characterized by a large change in the frequency of F_1 with an immobility of F_2. Monsen hypothesized that this pattern results when the appropriate jaw movement is not accompanied by appropriate movement of the tongue. Minimal movement of both F_1 and F_2 was another pattern observed, which Monsen attributed to a generally stable vocal tract throughout production of the diphthong with minimal jaw movement. A third pattern was a reversal of the direction of movement of F_2 with respect to normal. Monsen hypothesized this to be the acoustic consequence of the diphthong being produced with the tongue lowered and retracted.

IV. NON-SEGMENTAL PATTERNS IN THE SPEECH OF SEVERELY AND PROFOUNDLY HEARING-IMPAIRED SPEAKERS

This section will present information on the non-articulatory aspects of hearing-impaired children's speech. These patterns are also referred to as suprasegmental because they involve characteristics of speech that extend over units composed of more than one phonetic segment. Included in this category are characteristics such as timing, intonation, and stress assignments. These areas, as well as the acoustic correlate of pitch (fundamental frequency) and factors affecting perceived voice quality will be described in this section.

A. Timing Patterns

1. Overall Speaking Rate

With few exceptions, the speech of the severely and profoundly hearing impaired is perceived as being too slow and sounding very labored. Physical measures of speaking rate have shown that profoundly hearing-impaired speakers, on the average, take 1.5 to 2.0 times longer to produce the same utterance as do normal-hearing speakers (Boone, 1966; Heidinger, 1972; Hood, 1966; John & Howarth, 1965; Voelker, 1935, 1938). The reduced speaking rate is due to the excessive prolongation of speech segments and the insertion of pauses.

Prolongation of speech segments may be present in the production of phonemes, syllables, and words. Calvert (1961) was among the first to obtain objective measurements of phonemic duration in the speech of the hearing impaired by spectrographic analysis of bisyllabic words. The results of this study showed that hearing-impaired speakers extended the duration of vowels, fricatives, and the closure period of plosives up to five times the average duration for normal speakers. In a later study, Osberger and Levitt (1979) observed that syllable prolongation in the speech of the hearing impaired was due primarily to prolongation of vowels.

Figure 1 shows data obtained by Osberger (1978) on mean syllable duration in a sentence produced by six normally-hearing and six profoundly hearing-impaired children. The data show a distinctive pattern of syllable durations for the two groups of speakers. The line connecting the data points of the hearing-impaired speakers lies above and is approximately parallel to that of the hearing children. The exception to this is the sixth syllable where the mean syllable duration is shorter for the hearing-impaired than the normal speakers. This was due to the omission of some of the phonemes in the syllable by the hearing-impaired speakers, making the duration of the syllable shorter than would be expected if all of the intended segments had been produced. The size of the standard deviations, shown by the vertical bars, indicates that there is greater variability in syllable duration among the hearing-impaired than among the normal speakers.

Profoundly hearing-impaired speakers typically insert more pauses, and pauses of longer duration, than do speakers with normal hearing (Boone, 1966; Boothroyd, Nickerson, & Stevens, 1974; Heidinger, 1972; Hood, 1966; John & Howarth, 1965; Stevens, Nickerson, & Rollins, 1978). Pauses may be inserted at syntactically inappropriate boundaries such as between two syllables in a

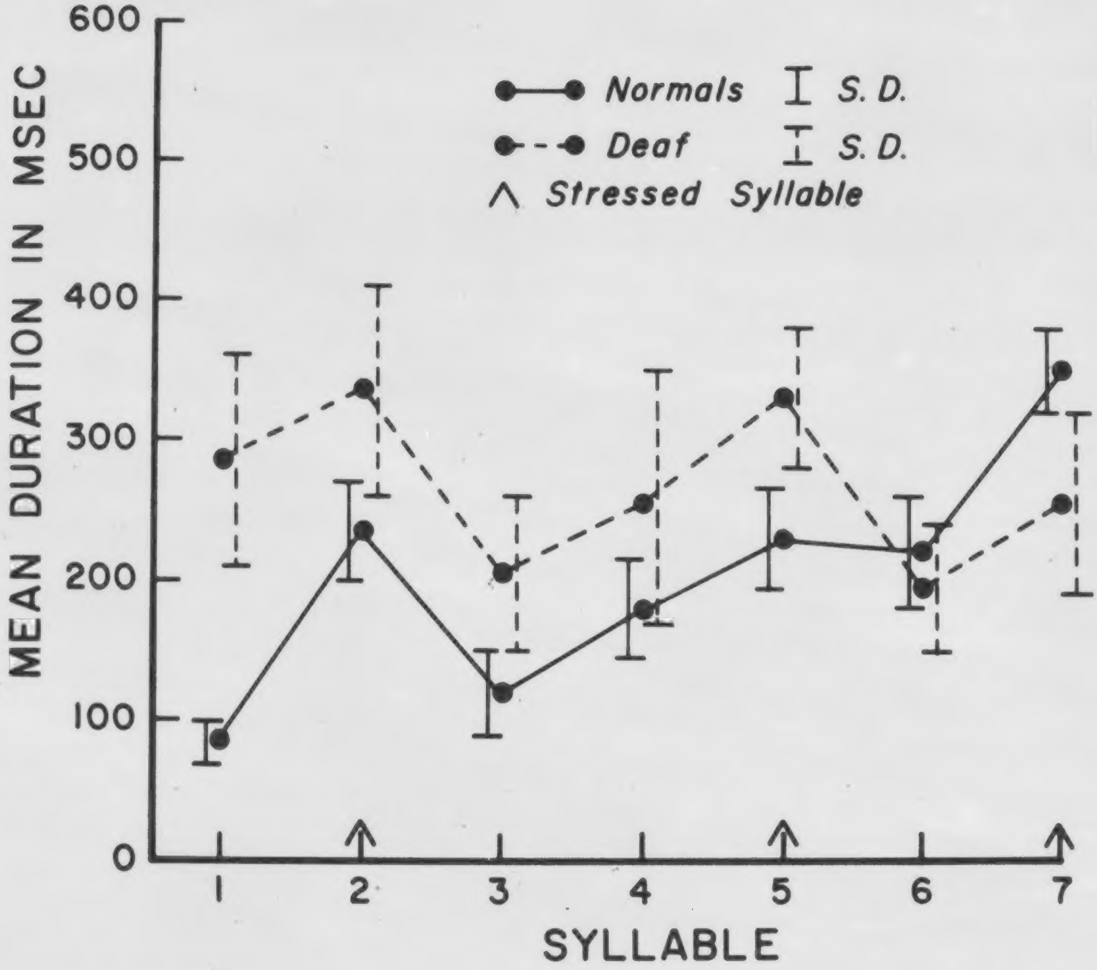


Figure 1. Mean duration (msec) for syllables in the sentence "I wish I could read that book" produced by six normal-hearing children and six hearing-impaired children. The standard deviation is represented by the vertical bars (after Osberger, 1978).

bisyllabic word or within phrases. The greatest difference between normal and hearing-impaired speakers has been observed in the durations of inter- and intraphrase pauses (Stevens et al., 1978). The results of Hudgins (1934, 1937, 1946) suggested that the frequent pauses observed in the speech of the hearing impaired may be the result of poor respiratory control. Specifically, Hudgins found that deaf children used short, irregular breath groups often with only one or two words, and breath pauses that interrupted the flow of speech at inappropriate places. In addition, there was excessive expenditure of breath on single syllables, false groupings of syllables, and misplacements of accents. Later, we shall discuss the propensity of hearing-impaired speakers to use inappropriate breath groups.

2. Segmental Timing Effects

Acoustic analyses of normal speech have shown that the duration of vowels is systematically influenced by effects operating at the level of phonetic segments. Since vowels form the nuclei of the larger segments of speech, these differences in vowel duration exert substantial effects on both the production and perception of the temporal and segmental aspects of speech. Vowels have been described as having an intrinsic duration (Peterson & Lehiste, 1960) and, in comparable contexts, some vowels are consistently shorter than other vowels (House, 1961). Hearing-impaired speakers with severe and profound losses have been found to distort this relationship between the vowels. For example, Monsen (1974) observed that /i/ was relatively longer than /I/ in monosyllabic words in the speech of normal-hearing subjects, but in the speech of profoundly hearing-impaired children, there was a tendency for these vowels to occupy mutually exclusive duration ranges. McGarr and Harris (1980), on the other hand, found that the profoundly hearing-impaired speaker in their study did not show consistent differences in intrinsic vowel duration.

There is substantial literature showing that the average duration of vowels also varies markedly as a function of phonetic context in normal speech. When different phonetic contexts are considered, the voicing characteristic of the following consonant has been shown to have a consistent effect on preceding vowel duration; for normal speakers, the duration of a vowel preceding a voiceless consonant is less than the vowel duration preceding a voiced consonant in stressed syllables (Denes, 1955; House, 1961; House & Fairbanks, 1953; Peterson & Lehiste, 1960). This systematic change in vowel duration has been found to be a significant perceptual cue to the voicing characteristic of the following consonant or consonant cluster (Raphael, 1972). Results obtained by Calvert (1961) and Monsen (1974) have shown that the hearing impaired fail to produce the appropriate modifications in vowel duration as a function of the voicing characteristics of the following consonant. Thus, the frequent voiced-voiceless confusions observed in the speech of the deaf may actually be due to vowel duration errors (Calvert, 1961).

3. Suprasegmental Timing Effects

The duration of segments is also influenced by effects operating at the level of syllables, words, and phrases. In English, changes in contrastive stress have been found to produce systematic changes in vowel duration. When vowels are stressed, they are longer in duration than when the same vowels are unstressed (Parmenter & Trevino, 1936). This durational variation has also been found to be an important cue for the perception of stress (Fry, 1955, 1958).

Several investigations have shown that while hearing-impaired speakers make the duration of unstressed syllables shorter than that of the stressed syllables, the proportional shortening is smaller, on the average, in the speech of the hearing impaired than in the speech of normal subjects (Osberger & Levitt, 1979; Stevens et al., 1978). In contrast to this, Reilly (1979) found larger than normal duration differences between vowels in primary- and weak-stress syllables produced by a group of profoundly hearing-impaired children. These data are shown in Figure 2. In this figure, duration has been calculated for the vowels /i, I, u/ produced in both primary- and weak-stress syllables by hearing and hearing-impaired children. For /i/ and /u/, longer average durations were measured for greater stress for both groups, with the hearing impaired durations being longer overall, and the difference between the primary and weak syllables being more extreme than in the samples produced by the hearing children. There was almost no difference in duration between the primary and weak /I/ in the normal children's samples, whereas the hearing-impaired speakers produced longer durations of /I/ in weak syllables than primary stress syllables.

Exactly how a hearing-impaired speaker uses temporal manipulations to convey differences in syllabic stress pattern is not clear. In a recent study, McGarr and Harris (1980) found that even though intended stressed vowels were always longer than unstressed vowels in the speech of one profoundly hearing-impaired speaker, the intended stress pattern was not always perceived correctly by a listener. Thus, the hearing-impaired speaker was using some other suprasegmental feature to convey contrastive stress. Variation in fundamental frequency would be a likely alternative, but McGarr and Harris also found that while the hearing-impaired speaker produced the systematic changes in fundamental frequency associated with syllable stress, perceptual confusions involving stress pattern were still observed.

Another suprasegmental temporal effect occurring in normal speech is prepausal lengthening. When a syllable occurs before a pause that marks a major syntactic boundary, it is longer in duration than when it occurs in other positions in a phrase (Klatt, 1975). It has been observed that hearing-impaired speakers do not always lengthen the duration of phrase-final syllables relative to the duration of the other syllables in the phrase. Stevens et al. (1978) observed that when there was evidence of prepausal lengthening in the speech of profoundly hearing-impaired talkers, the increase in the duration of the final syllable was much smaller, on the average, for the hearing-impaired than for the normal-hearing speakers. In contrast to this, Reilly (1979) found that the profoundly hearing-impaired speakers in her study used duration to differentiate prepausal and non-prepausal syllables. As was the case for primary- and weak-stress syllables, discussed above, Reilly observed a larger than normal difference between the duration of syllables in the prepausal and non-prepausal position in the samples produced by the hearing-impaired children.

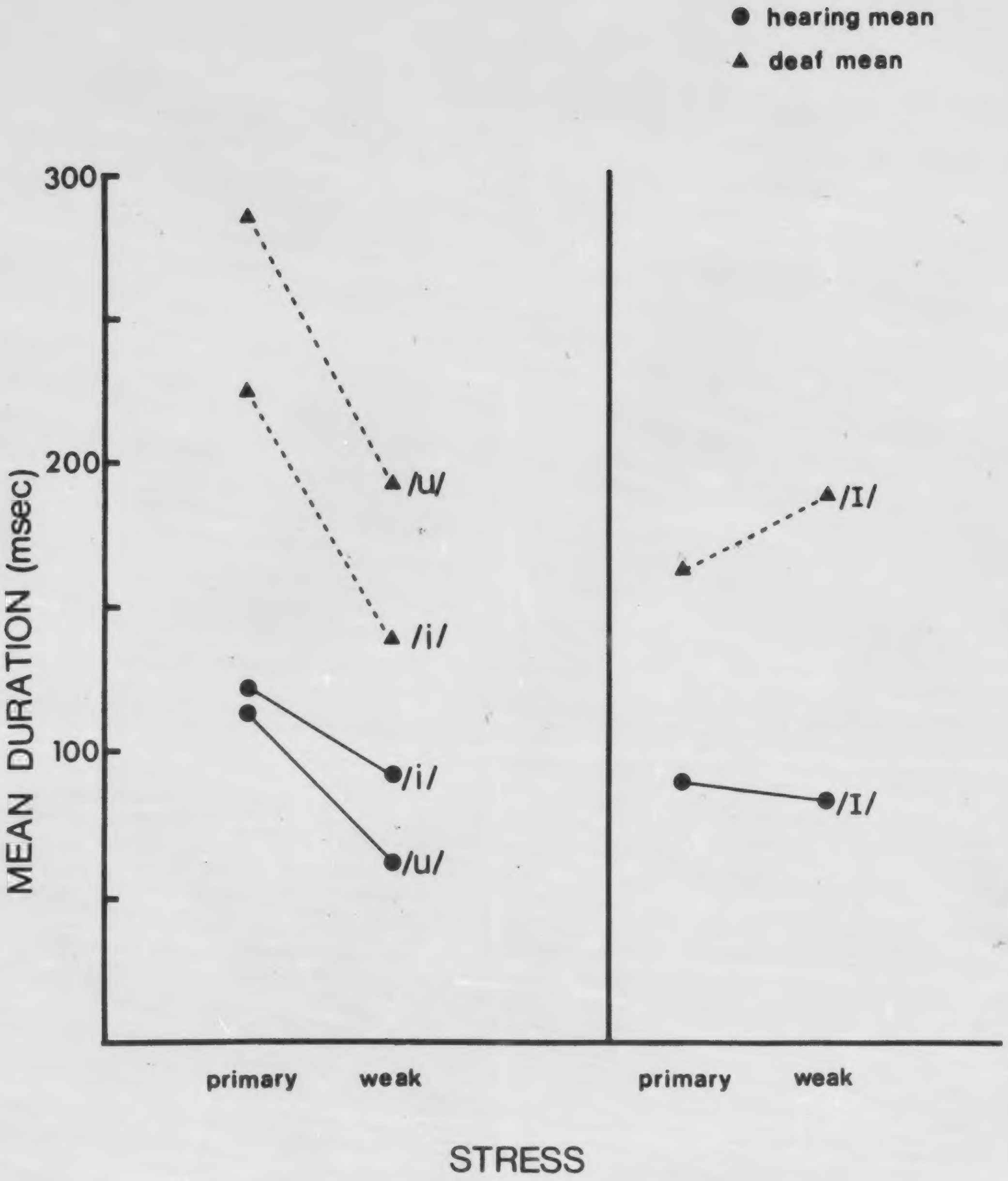


Figure 2. Mean vowel duration (msec) in primary- and weak-stress syllables produced by a group of normal-hearing and a group of profoundly hearing-impaired children (after Reilly, 1979).

The information presented above clearly shows that profoundly hearing-impaired speakers distort many temporal aspects of speech. These distortions, such as excessively prolonged speech segments, and the insertion of both frequent and lengthy pauses, are perceptually prominent and disrupt the rhythmic aspects of speech. In spite of these deviancies, there is evidence suggesting that hearing-impaired talkers manipulate some aspects of duration, such as those involving relative duration, in a manner similar to that of speakers with normal hearing.

B. Fundamental Frequency Patterns

1. Average Fundamental Frequency

Among the most noticeable speech disorders of the hearing impaired are those involving fundamental frequency (Fo). In normal speech, there are differences in average fundamental frequency depending on the sex and age of the speaker. Reported fundamental frequency values range from 100-175 Hz for adult males and from 175-250 Hz for adult females (Fairbanks, 1940; Fairbanks, Wiley, & Lassman, 1949; Fairbanks, Herbert, & Hammond, 1949; Hollien & Paul, 1969). Recent data (Hasek, Singh, & Murry, 1981) suggest that a significant difference between the average Fo of preadolescent male and female children with normal hearing begins to emerge by seven or eight years of age, with the sex difference attributable to a reduction in Fo for male children only, beginning around age seven. No significant preadolescent age-related change in Fo in females was observed.

If there is a problem with a hearing-impaired speaker's average fundamental frequency, more often the voice pitch is characterized as too high rather than too low (Angelocci et al., 1964; Boone, 1966; Martony, 1968). Some differences in average Fo have been found as a function of the age or sex of the hearing-impaired speaker. The results of several studies have shown that there are no significant differences in average Fo between young hearing and hearing-impaired children in the 6-12 year age range (Boone, 1966; Green, 1956; Monsen, 1979). Differences have been reported between groups of older children but it is not clear if pitch deviation is greater for hearing-impaired females or males. Boone (1966) found a higher average Fo for 17-18 year old males than females. Osberger (1981) found that the difference in Fo between hearing and hearing-impaired speakers in the 13-15 year age range was greater for females than for males. This finding is illustrated in Figure 3, which shows the Fo values averaged across sentences for six normal-hearing and ten hearing-impaired subjects. As can be seen, the Fo for the female hearing-impaired speakers ranged between 250-300 Hz. This value is about 75 Hz higher than that observed for the normal-hearing females. The average Fo value of the utterances of the male hearing-impaired speakers is slightly lower than that of the hearing males for the first part of the utterance. The Fo values for the hearing and hearing-impaired male speakers overlap for the last half of the utterance. Bush (1981) observed excessive segmental variations in Fo for a small group of profoundly hearing-impaired females in the same age range as those in the Osberger study. Age-related factors such as laryngeal growth accompanied by adolescent voice change or similarities in speech training were suggested by Bush as reasons for the problems of the females in controlling Fo.

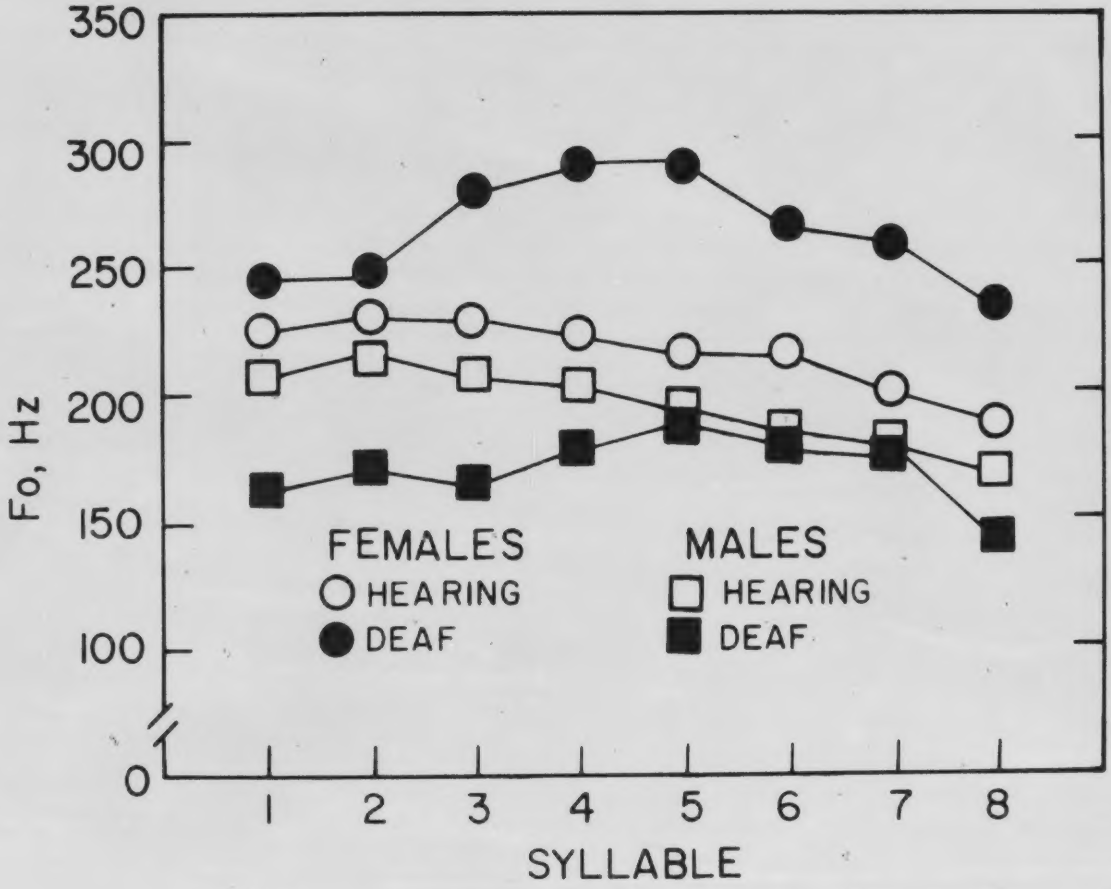


Figure 3. Fundamental frequency values in Hz, measured at the center of the vowel in each syllable in the sentence "I like happy movies better" for groups of normal-hearing and profoundly hearing-impaired males and females.

Up to this point, we have limited our discussion to physical measures of fundamental frequency. In a clinical or school situation, the examiner will not, in most cases, have the equipment necessary to make such measurements. In these settings, the clinician will have to rely on his or her perceptual abilities to evaluate the appropriateness of the child's pitch. The pitch deviancy of profoundly hearing-impaired children has been evaluated perceptually by McGarr and Osberger (1978), using a five-point rating scale. The profile rating of pitch register (Subtelny, 1975) and the descriptors are shown in Table 1. The scale was used with approximately 50 children 10-11 years of age. The results of this study showed that a large number of the children received pitch ratings that were either appropriate for their age and sex or differed only slightly from optimal level. Thirty-two of the children received an average rating higher than 4.0. There was, however, a small group of children who could not sustain phonation and whose speech was characterized by pitch breaks or large fluctuations in pitch. On the whole, these findings are in agreement with earlier studies indicating that the pitch of many preadolescent hearing-impaired children is within the normal range. It is not clear to what extent the average Fo of a hearing-impaired child's speech can differ from that of a normal child before it is perceived as deviant and hence, remedial training is indicated.

Table 1

Rating Scale Used to Evaluate Pitch (from Subtelny, 1975)

Profile Rating	Functional Descriptor
1.	Cannot sustain phonation
2.	Much above (+) or much below (-) optimal level
3.	Moderately above (+) or below (-) optimal level
4.	Slightly above (+) or below (-) optimal level
5.	Appropriate for age and sex

2. Intonation Patterns

Intonation is the perceived pattern of change in fundamental frequency within a phrase or sentence. Reference is made, even in the very early literature, to the difficulties that hearing-impaired speakers experience in controlling this aspect of speech. Haycock (1933), Rawlings (1935), Russell (1929), Scripture (1913), and Story (1917) all describe the speech of congenitally deaf persons as "monotonous" and "devoid of melody." Later investigations showed that hearing-impaired speakers did produce pitch variations, but the average range was more reduced than those of speakers with normal hearing (Green, 1956; Hood, 1966; Voelker, 1935).

Some hearing-impaired speakers may demonstrate an intonation problem in the form of excessive and inappropriate changes in fundamental frequency. These speakers may raise or lower Fo 100 Hz or more within the same utterance. Often, after a sharp rise in fundamental frequency, the hearing-impaired speaker loses all phonatory control and there is a complete cessation of phonation (Monsen, 1979; Smith, 1975; Stevens et al., 1978).

Figure 4 shows the intonation contour of a simple, declarative sentence spoken by a normal, 14-year-old female. There is a rise in Fo at the beginning of the sentence with a peak on the first stressed syllable (the second syllable in the sentence). As the sentence is produced, there is a gradual reduction in Fo, known as declination. The sharp drop that occurs in Fo at the end of the sentence is referred to as the terminal fall. Figure 5 shows the contour of the same sentence spoken by a hearing-impaired male speaker, 14 years of age, judged to have insufficient variation in intonation. Note that the extent of the change in the Fo throughout the utterance is more restricted than that observed for the child with normal speech in Figure 4. In contrast to this pattern, Figure 6 shows contours for two females, 14 years old, who produced the sentence with excessive and inappropriate changes in Fo. Speaker 1 produced the first part of the sentence with a sharp rise in Fo, followed by a sharp fall in Fo over the last half of the utterance. Speaker 2 produced inappropriate fluctuations in Fo throughout the entire utterance.

There have been few attempts to arrive at a quantitative classification of intonation contours produced by hearing-impaired children. Monsen (1979) has described the following four types of contours that he found to occur in the production of CV syllables by 3- to 6-year-old hearing-impaired children: (1) a falling contour, characterized by a smooth decline in Fo at an average rate greater than 10 Hz per 100 msec; (2) a short-falling contour, occurring on words of short duration--the Fo fall may be more than 10 Hz per 100 msec but the total change may be small; (3) a falling-flat contour, characterized by a rapid change in frequency at the beginning of a word, followed by a relatively unchanging, flat portion; (4) a changing contour, characterized by a change in frequency, the duration of which appears uncontrolled, and extends over relatively large segments.

Monsen (1979) found that the type of contour appeared to be an important characteristic in separating the better from the poorer hearing-impaired speakers. His classification scheme represents a substantial step forward in describing the intonation patterns of the hearing impaired. It remains to be determined if such a classification scheme can be used to describe objectively the intonation patterns of entire sentences as well as isolated syllables.

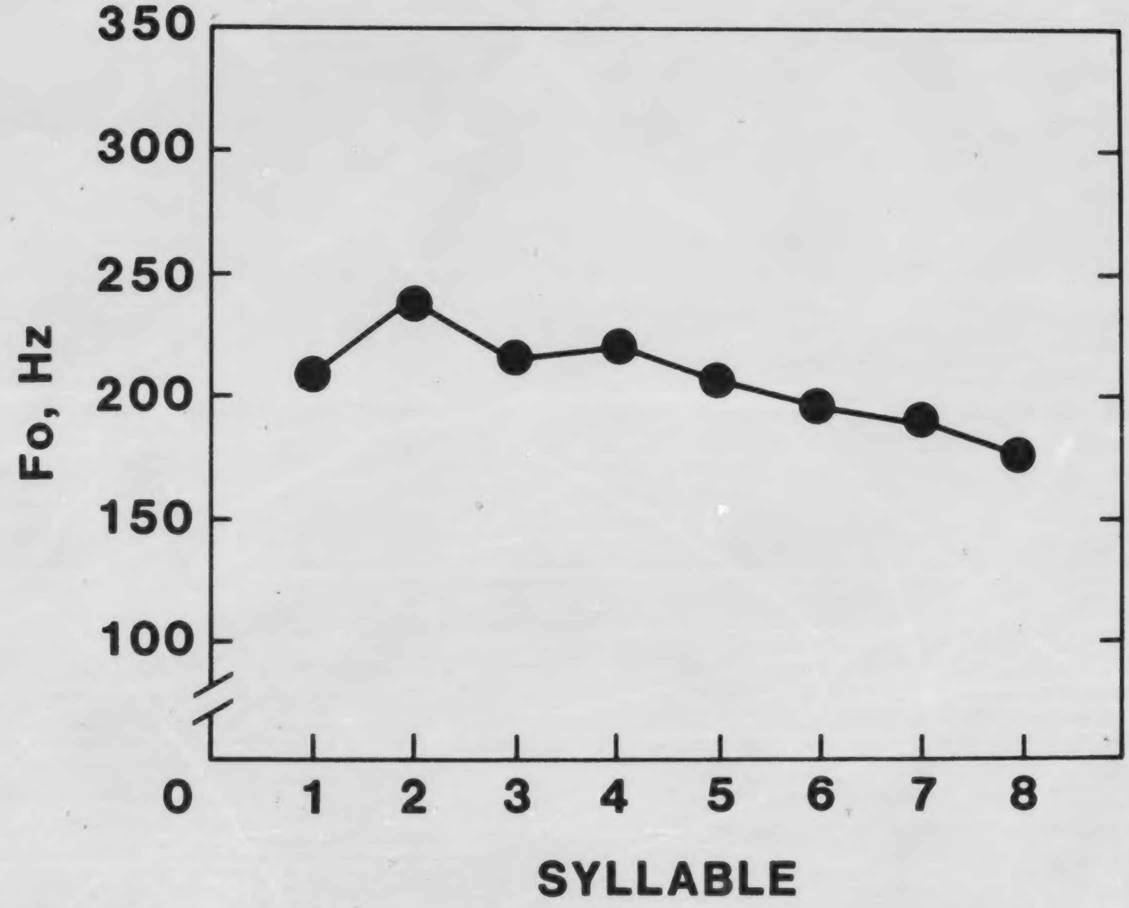


Figure 4. The intonation contour of the simple declarative sentence, "I like happy movies better," spoken by a normal-hearing child. Each data point is the fundamental frequency value in Hz, measured at the center of the vowel in each syllable in the sentence.

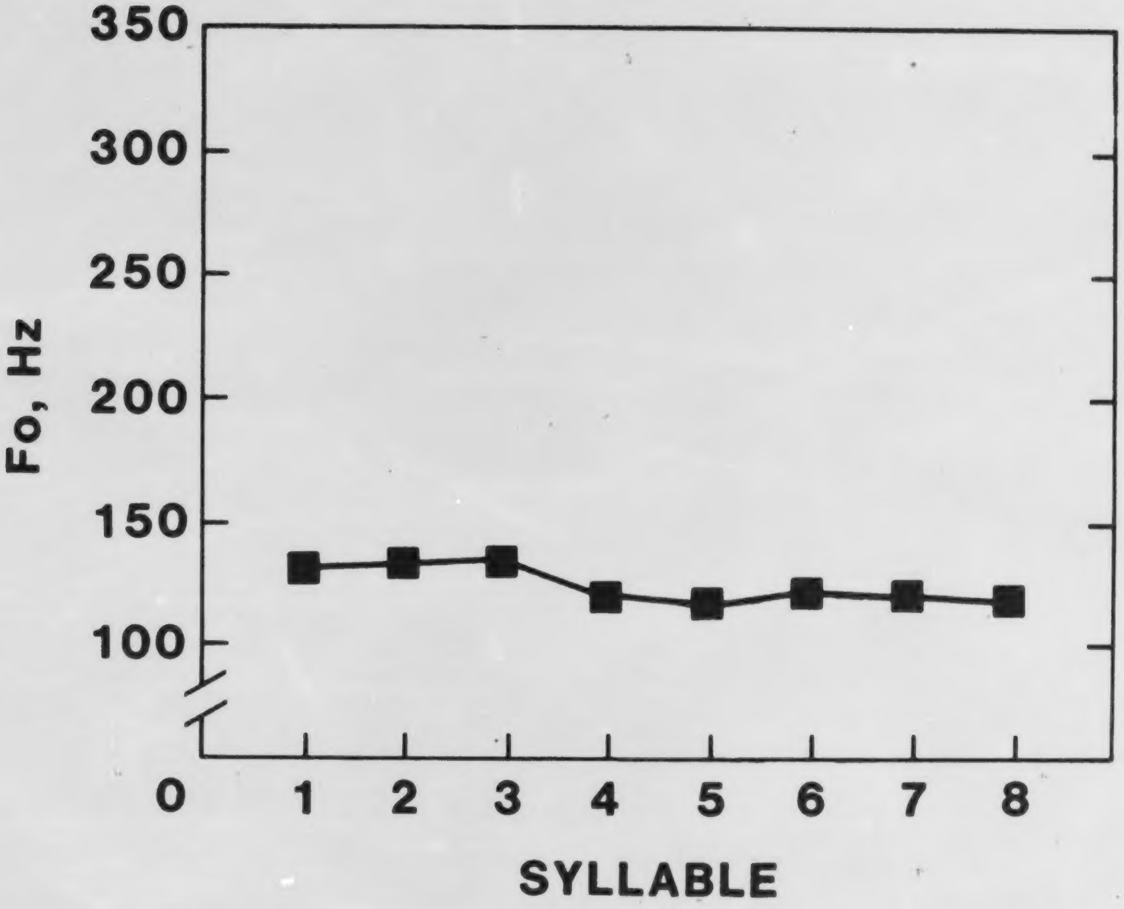


Figure 5. The intonation contour of the sentence "I like happy movies better" spoken by a profoundly hearing-impaired speaker judged to produce insufficient variation in intonation. Each data point is the fundamental frequency value in Hz, measured at the center of the vowel in each syllable of the sentence.

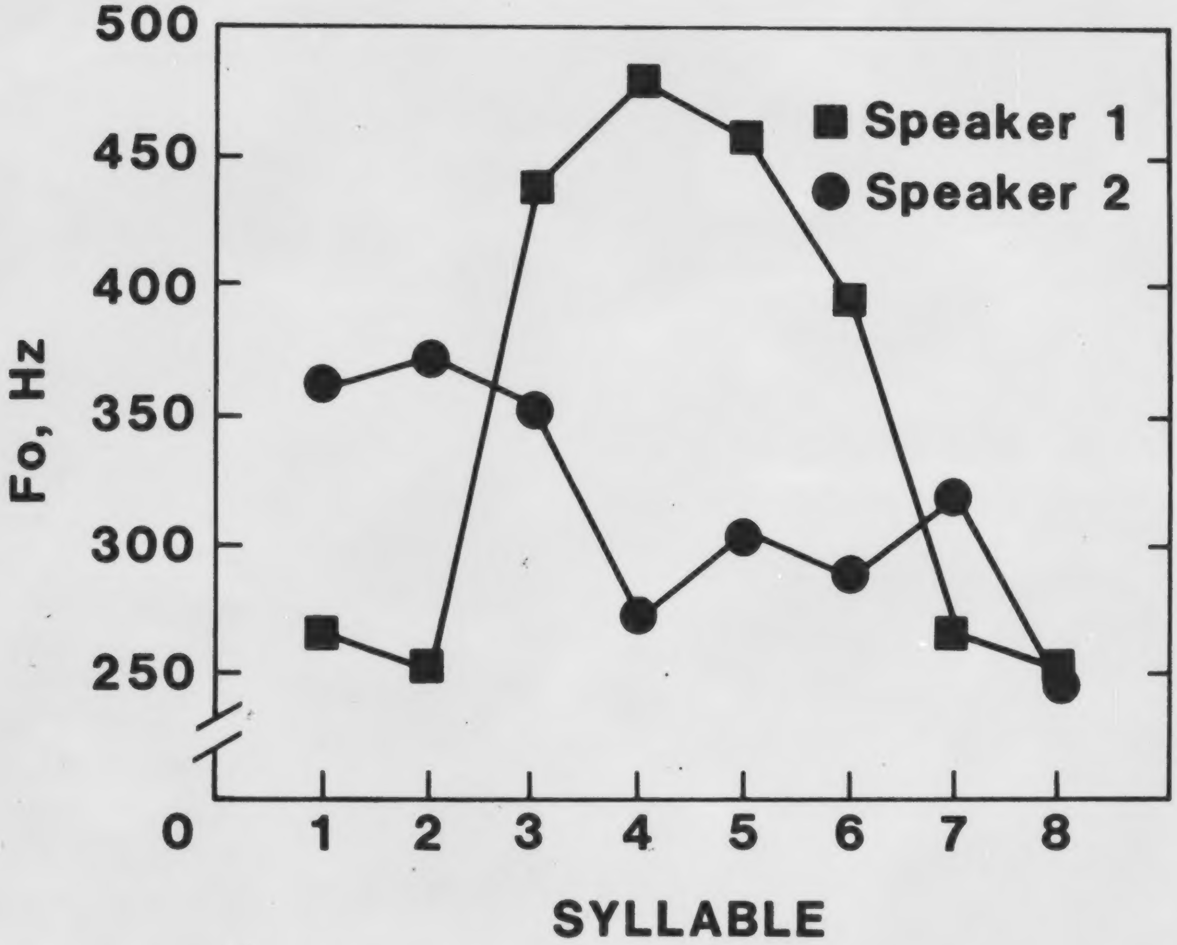


Figure 6. The intonation contour of the sentence "I like happy movies better" spoken by two profoundly hearing-impaired females judged to produce excessive and inappropriate changes in fundamental frequency. Each data point is the fundamental frequency value in Hz, measured at the center of the vowel in each syllable of the sentence.

One factor that strongly influences Fo changes is the degree of stress placed on syllables within a breath group. Typically, stressed syllables are spoken with a higher fundamental frequency than are unstressed syllables (Fry, 1955). Thus, the contour consists of peaks (rises) and valleys (falls) in Fo that correspond to the stressed and unstressed syllable pattern of the sentence. This pattern has been observed to be distorted in the speech of the hearing impaired. An example of this distortion is apparent in the Fo contours of the two speakers in Figure 6.

3. Segmental Influences on Fundamental Frequency Control

A common clinical observation is that some hearing-impaired children produce the vowels /i, I, u/ on a higher Fo than the other vowels of English. It has been shown that there is a systematic relationship between vowel and Fo in normal speech. High vowels are produced on a higher Fo than low vowels, resulting in an inverse relationship between Fo and the frequency location of the first formant of the vowel (House & Fairbanks, 1953; Peterson & Barney, 1952). Angelocci et al. (1964) first examined some of the vowel changes in Fo in the speech of the hearing impaired. Their results showed that the average Fo and amplitude for all vowels was considerably higher for the hearing-impaired than for the normal subjects. In contrast, the range of frequency and amplitude values for the vowel formants was greater for the normal-hearing than for the hearing-impaired speakers. This finding, combined with the high Fo and large amplitude values, led Angelocci et al. to suggest that the hearing-impaired subjects attempted to differentiate vowels by excessive laryngeal variation rather than with articulatory maneuvers as do normal-hearing speakers.

A recent study by Bush (1981) does not support a simple trade-off between Fo variability and articulatory skill. Bush observed a close relationship between vowel-related variability in Fo and articulatory skill for the majority of profoundly hearing-impaired subjects in her study. In general, greater Fo variability was observed for the hearing-impaired speakers who produced a wide range of vowel sounds (in terms of F_1 and F_2 values) and who were more intelligible than speakers whose articulatory skills were more limited. Bush also noted that although the amount of Fo variation with vowels used by the hearing-impaired speakers was greater, on the average, than that used by the hearing speakers, the direction in which Fo varied as a function of vowel height was similar for the two groups of speakers.

On the basis of these observations, Bush concluded that the vowel-to-vowel variations produced by the hearing-impaired speakers were, in some way, a consequence of the same articulatory maneuver used by normal speakers in vowel production. These data are discussed in terms of Honda's (Note 1) account of vowel-related variations in Fo for normal speakers. Briefly, Honda's mechanism assumed that moving the tongue root forward for the production of high vowels causes the thyroid bone to move forward, tilting the cartilage anteriorly. As a result of these maneuvers, there is increased tension on the vocal folds, resulting in an increase in Fo. Bush has postulated that because of the non-linear nature of the stress-strain relationship for vocal-fold tissue, increases in vocal-fold tension may be greater in magnitude when the tension on the vocal folds is already relatively high (as is the case with hearing-impaired speakers), resulting in somewhat larger increases in Fo during the articulation of high vowels.

In summary, as was observed for some of the temporal patterns of speech, it appears that profoundly hearing-impaired speakers encode and organize some aspects of fundamental frequency with respect to syntactic considerations in much the same manner as do normal speakers. There are obvious deviancies in fundamental frequency control in the speech of the hearing impaired, but there is evidence to suggest that they know and use some of the same rules applied by normal-hearing speakers.

C. Production Patterns Affecting Voice Quality

1. Voice Quality

It is not unusual to find people who, after working with the profoundly hearing impaired, claim that the speech of this population has a distinctive quality that differentiates it from other speakers. Calvert (1961) found that teachers of the hearing impaired could reliably differentiate the voices of profoundly hearing-impaired speakers from normal speakers, provided the speech samples contained articulatory movement, such as that required for the production of a diphthong or a CVC syllable. Productions with negligible articulatory movements, such as sustained vowels, failed to provide the experienced listeners with the necessary information for the correct identification of speakers. On the basis of these findings, Calvert concluded that the distinguishing characteristics of the speech of the profoundly hearing impaired are associated with articulatory movement over time, rather than with voice quality per se.

In the same study, Calvert (1961) also found that there was a great deal of variability among teachers in choosing the characteristics they felt described most closely the voice quality of the hearing impaired. Thus, although the deviant voice quality of the hearing impaired can be recognized easily, the characteristics that contribute to the perceived deviation are difficult to characterize.

In a recent study, Monsen (1979) quantified some of these characteristics. Acoustic analysis of duration, fundamental frequency, and phonatory control were correlated with ratings of voice quality for monosyllables produced by young hearing-impaired children. The results of this study showed that the fundamental frequency contour appeared to be the most general acoustic characteristic differentiating the children with better voices from those with poorer voices. Children with good voice-quality ratings had fundamental frequency contours that fell within an appropriate range and that varied over time in an appropriate manner. In contrast, children with poor voice quality produced intonation contours that were excessively flat or excessively changing. Monsen (1979) concluded that while other deviations such as poor vowel quality, breathiness, and duration errors may exert a strong influence on perceived voice quality in individual cases, they do not appear to be the major factors in determining the quality of the voice. From the results of this study and those of Calvert (1961), it appears that the distinctive voice quality of the hearing impaired may be due to both poor articulatory timing control and inadequate control of source function.

2. Nasalization

Proper control of the velopharynx has been recognized as a source of difficulty for hearing-impaired speakers for many years (Hudgins, 1934). If the velopharyngeal port is opened when it should be closed, the speech may be perceived as hypernasal; if it is closed when it should be opened, hyponasality will result. Problems in nasalization control are often described as affecting voice quality because hyper- or hyponasality affects the resonant properties of speech. Improper velopharyngeal control may also result in articulatory errors, a problem addressed earlier in this chapter.

In a clinical setting, the evaluation of velopharyngeal control is usually made on the basis of qualitative judgments, which are often difficult to assess because they may be influenced by the presence of other deviancies. Stevens et al. (1976) have attempted to overcome this problem by developing a procedure to quantify the degree of nasalization for nasal and non-nasal sounds in the speech of hearing-impaired children. Measurements of nasalization have been obtained with an accelerometer attached to the surface of the nose. The accelerometer picks up vibrations of the nose when there is velopharyngeal opening during a voiced sound. Stevens et al. have evaluated adequacy of velar control by comparing the amplitude of the accelerometer signal (in decibels) for nasal consonants to the amplitude of vowel sounds that should be produced without nasalization. For normal-hearing speakers, the amplitude difference between these measures is in the range of 10-20 dB. Using amplitude difference as an index of nasalization, Stevens et al. found that 76% of the profoundly hearing-impaired children studied had excessive nasalization in at least half of the vowels produced in monosyllabic words. Excessive nasalization on at least 8 of the 10 vowels studied was observed for 36% of the children. The greatest difficulty in velopharyngeal control was evidenced in the hearing-impaired children's production of nasal-stop clusters, which required closely coordinated movements of the velopharynx and oral articulators. Almost half of the hearing-impaired children made an error on at least one word with a nasal-stop cluster.

3. Breathy Voice and Glottalization

These problems are caused by improper adjustment of the vocal folds. Breathiness occurs when there is excessive airflow during voicing, resulting in generation of turbulence noise at the glottis. The vocal folds do not come together rapidly, which affects the shape of the volume-velocity waveform, resulting in an acoustic waveform with enhanced energy in the low frequencies and deficient energy in the high frequencies (Stevens et al., 1978).

Glottalization involves the insertion of the glottal stop between syllables or words. It is caused by tightly adducting the glottal folds and then abruptly releasing them. Profoundly hearing-impaired children often substitute glottal stops for consonants produced in the center and back of the mouth (Levitt et al., 1976). There is a tendency for hearing-impaired children who insert many glottalizations in their speech to have lower intelligibility than those who do not (Stevens et al., 1978).

V. PRODUCTION PATTERNS IN THE SPEECH OF HARD-OF-HEARING CHILDREN

A. Articulatory Patterns

Until only recently, little attention has been paid to the speech of the hard-of-hearing child. This is probably largely due to the fact that the majority of these children are integrated into regular schools and they are not as accessible for study as the students attending day schools for the deaf. In addition, researchers traditionally have viewed the communication and education problems of the profoundly hearing impaired as more serious than those of the hard of hearing and, thus, the majority of research effort has been devoted to the children who appeared to have the greatest need. We now know that the presence of even a mild hearing loss can affect speech and language development and interfere with academic performance. Often, hard-of-hearing children are neglected in the public school system. They frequently fail to receive the support services from appropriately trained professionals that they require in order to perform successfully in a regular class (Davis, 1977).

The majority of information available on the speech of hard-of-hearing children involves analyses of articulatory skills. Relatively few studies have quantified suprasegmental production patterns and, for this reason, only the segmental aspects of the speech of hard-of-hearing children will be discussed.

If it is assumed that the major difference between hard-of-hearing and profoundly hearing-impaired children is the degree of hearing loss, it is to be expected that hard-of-hearing children would have better speech skills than children with profound hearing losses. This notion has, in fact, been supported by the results of several studies showing that, on the average, there is a lower frequency of vowel and consonant errors in the speech of hard-of-hearing children than in the speech of profoundly hearing-impaired children (Gold, 1978; Hudgins & Numbers, 1942; Markides, 1970; Nober, 1967).

Probably the most comprehensive study on the speech of hard-of-hearing children has been completed by Gold (1978). In this study, the articulatory errors made by mainstreamed heard-of-hearing (Pure Tone Average of 80 dB HTL or less) and deaf (PTA of 80 dB HTL or greater) children were compared. Phonemic transcriptions were made of sentences read by the children that contained all the phonemes of English. The data were analyzed to determine if the types of articulatory errors were the same for the two groups of children. The results in terms of overall error rate revealed, not unexpectedly, that the deaf group had significantly more segmental errors than the hard-of-hearing group. The data further revealed that the types of errors were similar for the two groups of children. These data are summarized in Table 2. Two calculations were made for each of the eight error types for both groups of children. The first calculation, error type as the proportion of intended phonemes for each of the two groups (shown in the first column of Table 2), was derived from the frequency of the error type relative to the total number of phonemes in the sample. The second calculation, error type as a proportion of all of the errors (shown in the second column), was performed to take into account the higher error rate of the deaf group. Thus, the proportion of

---

Table 2

Relative Frequency of Articulatory Errors for Hard-of Hearing and Deaf Children (from Gold, 1978).

	Hard-of-Hearing		Deaf	
	Proportion of Intended Phonemes	Proportion of Errors	Proportion of Intended Phonemes	Proportion of Errors
Type of Error				
Omissions	.076	(.392)	.116	(.405)
Vowel-Vowel Substitutions	.050	(.258)	.065	(.227)
Consonant-Consonant Substitutions	.035	(.180)	.060	(.210)
Recognizable Distortions	.019	(.098)	.023	(.080)
Unrecognizable Distortions	.007	(.036)	.013	(.045)
Non-English Substitutions	.002	(.010)	.004	(.014)
Diphthong Errors	.004	(.021)	.004	(.014)
Other	.001	(.005)	.001	(.007)
Total Proportion of Error	.194	(1.000)	.286	(1.000)

---

errors was based on the relative frequency of the error type out of the total proportion of errors made by the group. Once the overall rate was taken into account, the data showed striking similarities in the frequency of an error type for the hard-of-hearing and deaf children. For example, the most frequent type of error for both groups was that of omission. As the data show, differences between the two groups were not substantial.

The results of the Gold study showed that although profoundly hearing-impaired children produce more segmental errors than hard-of-hearing children, the relative proportion of errors for each error type is similar for both groups of children. Only a small number of phonemes showed any significant differences in the pattern of confusions between groups. Gold has concluded, at least for children in the same type of educational setting, that the degree of the hearing loss is more strongly related to the overall frequency of errors than to the kinds of errors that will be made.

B. Pattern of Speech Errors of Different Populations of Children

From the preceding discussion, it becomes evident that the pattern of articulatory errors is remarkably similar in the speech of different populations of hearing-impaired children. Two studies, those of Smith (1975) and Gold (1978), lend themselves to cross-population comparison because the same test materials and procedures were used by the two investigators. The major difference between the studies is the groups of children studied: Smith examined the segmental errors in the speech of profoundly hearing-impaired children in an oral day school for the deaf; Gold, as mentioned above, examined the segmental errors in mainstreamed hard-of-hearing and profoundly hearing-impaired children. Some of the data from these two studies have been plotted in Figure 7. In this figure, correct production of consonants is plotted as a function of place of production for the three groups of hearing-impaired children. Comparison of the data shows distinct patterns across groups of children. As might be expected, the hard-of-hearing children most often produced the consonants correctly, followed by the mainstreamed profoundly hearing-impaired children; the children in the school for the deaf produced the consonants correctly the least often. Note also that sounds produced in the front of the mouth were most often correct, followed by the back sounds; sounds produced in the middle of the mouth were most prone to error, a finding discussed earlier in this chapter.

Gold did find some significant differences in the pattern of confusions made by the mainstreamed hearing-impaired children and the children in the school for the deaf. The children in the school for the deaf used more neutral vowel substitutions and omitted more consonants than did the mainstreamed children. They also substituted the glottal stop for /t/ and /k/, and /b/ for labial sounds more often than the profoundly hearing-impiared children who were mainstreamed.

The results of Gold's (1978) study show that although the nature of the confusions did not differ significantly between the hard-of-hearing and deaf children in the same educational setting, there were significant differences between the deaf children in schools for the deaf and those in the regular public schools. Similarities in segmental error pattern were also apparent across groups of children. It should be mentioned that although the

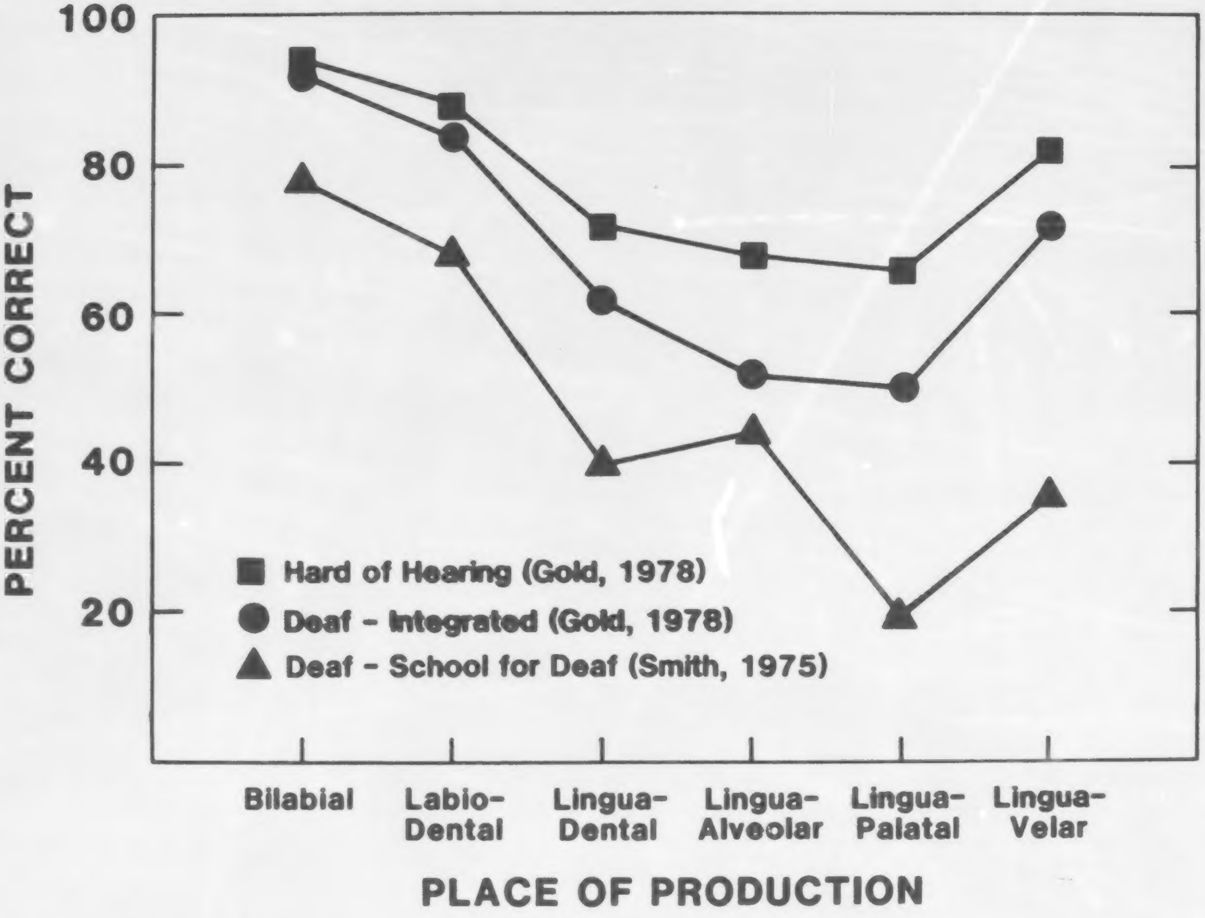


Figure 7. Percent correct production of consonants plotted as a function of place of production for three groups of hearing-impaired children (after Smith, 1975 and Gold, 1978).

mainstreamed children had better speech skills than the children in the school for the deaf, a causal relationship between speech skills and school setting cannot be concluded. Although it is possible that a hearing-impiared child's speech may improve as a result of daily exposure to hearing children, the children in Gold's study may have been mainstreamed because of their good speech skills.

VI. MECHANISMS OF PRODUCTION CONTROL

As we have described above, speech production skills of the hearing impaired have been examined using listener judgments, phonetic transcriptions, and acoustic analyses. While the descriptive literature is fairly detailed, there have been few physiological studies on the speech of the hearing impaired. This is surprising since technology is available and also because speech production skills in normals have been studied fairly extensively. Indeed, close to 50 years ago, researchers attempted objective measurements of hearing-impaired speech production in such areas as breath control (Hudgins, 1936; Rawlings, 1935; Scuri, 1935), voice production (Hudgins, 1937; Voelker, 1938), and articulation (Brehm, 1922; Hudgins, 1934). Although by today's technological standards the instrumentation in these studies was not very sophisticated, these researchers deserve our admiration for their ingenuity and creative insight. Their intuition and observations are clearly not dated. Consider the following:

> "The most obvious fault in the speech-breathing of deaf children is that they have little or no control over the breath supply so that a great deal more breath than is necessary is allowed to escape on each syllable. They do not speak with normal chest-abdominal action. They have not learned to group their syllables into breath groups and phrases. Instead, they often expend an entire breath on a single word. The reasons for this excessive use of breath is two-fold: The inco-ordinated (sic) movements of the breathing muscles and the maladjusted glottis" (Hudgins, 1937, p. 348).

The observations of Hudgins and his contemporaries might be taken today as evidence for a breakdown in interarticulator coordination. That is, hearing-impaired speakers fail to coordinate the complex activity of respiration, phonation, and articulation, and the resultant errors in timing occur at the segmental and suprasegmental levels of speech production.

Admittedly, there has been a long hiatus between the early research efforts and contemporary rekindled interest in speech physiology of the deaf. Whether the time lapse represents a period of preoccupied interest with that of describing the error patterns of hearing-impaired talkers, or reflects a lag in applying the technology and ideas of speech production in normals to speech production of the hearing impaired can only be conjecture. There may be some truth in each, but in any event, we now turn to some recent studies on the physiological characteristics of deaf speech.

A. Respiration

Studies on the respiratory patterns of profoundly hearing-impaired speakers have shown that they evidence at least two kinds of problems. The first is that they initiate phonation at too low a level of vital capacity, and also that they produce a reduced number of syllables per breath (Forner & Hixon, 1977; Whitehead, in press). The second problem is that they mismanage the volume of air by inappropriate valving at the laryngeal level.

Recently, Hixon and his associates have provided some objective data on respiratory behavior both in normal (Hixon, Goldman, & Mead, 1973; Hixon, Mead, & Goldman, 1976) and also in hearing-impaired speakers (Forner & Hixon, 1977). In these studies, magnetometers were used to measure changes in the anterior-posterior dimensions of the chest wall during respiratory maneuvers and also speech. Hearing-impaired speakers were found to be like hearing speakers in some respects but not in others. For example, respiratory activity for non-speech activities such as tidal breathing was similar to normal. This has also been noted for other non-speech respiratory activities such as coordinative demands on the breathing mechanism for athletics. Also, Forner and Hixon (1977) showed that the mechanical adjustments of the respiratory mechanism in preparing to speak (i.e., the relative posture of the rib cage versus the abdomen) were often correct. These findings do not support the suggestions of the early researchers that hearing-impaired speakers evidence inappropriate posture problems such as rib cage-abdominal asynchronization. However, like the early researchers, Forner and Hixon reported that hearing-impaired speakers paused at inappropriate linguistic boundaries either to inspire or alternatively to waste air, and thus they produced fewer syllables per breath unit. Hearing-impaired speakers were also found to initiate phonation at inappropriate lung volumes and to speak within a fairly restricted lung volume range.

These results have been confirmed by Whitehead (in press), who has extended the findings of Hixon by examining different respiratory patterns with respect to the speech intelligibility of hearing-impaired talkers. Not surprisingly, Whitehead showed that profoundly hearing-impaired speakers who were more intelligible had respiratory patterns similar to those of normal speakers. For example, both groups initiated speech well above functional residual capacity (FRC) and terminated production within the mid-volume range. In contrast, hearing-impaired speakers who were characterized as semi-intelligible initiated speech at substantially lower lung volumes and continued speaking well below FRC. Speech attempted at such reduced lung volumes is exceedingly difficult because the speaker is working against the natural recoil forces of the respiratory mechanism. Furthermore, this aberrant respiratory pattern will also directly affect phonation.

Control of the expiratory cycle for speech is crucial for phonation and is particularly important in producing events such as changes in vocal intensity, accommodating different aerodynamic patterns associated with consonant production, as well as linguistic phrasing. To achieve such speech events, the volume of expired air must be appropriately managed, and this usually occurs at the laryngeal level. Thus, during speech production, one might think of the relationship of the larynx to the respiratory mechanism as analogous to that of an air valve, where the valve must be open at certain times to let the air escape (e.g., when producing a voiceless segment), and must be closed at the other times to preserve the breathstream.

There are data suggesting that hearing-impaired speakers have difficulty in coordinating the events of respiration and laryngeal valving. For example, consider some aerodynamic studies of consonants produced by hearing-impaired speakers (Hutchinson & Smith, 1976; Whitehead, in press; Whitehead & Barefoot, 1980). The method of data collection in these studies was similar; air flow was measured using a face mask coupled to a pneumotachograph. Air flow measurements are taken to reflect the relative open or closed state of the vocal tract. For normals, voiceless plosives would be produced with greater peak airflow than their voiced cognates; fricatives would be produced with greater airflow than plosives. Overall, Whitehead and others cited above have shown that hearing-impaired speakers do produce, although inconsistently, plosives and fricatives with normal airflow patterns, suggesting that at least some hearing-impaired speakers are relatively successful in coordinating respiration and laryngeal valving. Not surprising, these speakers were among the more intelligible in the Whitehead study. Less intelligible hearing-impaired speakers were often quite variable in management of airflow, and they did not differentiate voiced and voiceless cognates aerodynamically. Data from these subjects suggest inappropriate laryngeal gestures that could reduce airflow or, in other words, an inability of some hearing-impaired speakers to coordinate respiration and laryngeal valving.

Another example of laryngeal valving problems can be gleaned from a study of laryngeal-supralaryngeal coordination in the speech of the hearing impaired (McGarr & Löfqvist, in press). In this experiment, laryngeal activity was monitored by means of transillumination, wherein a flexible fiberscope is used to illuminate the larynx, and a phototransistor, placed on the surface of the subject's neck below the cricoid cartilage, senses the light passing between the vocal folds. Figure 8 shows selected tokens of an utterance produced by one profoundly hearing-impaired speaker. Information about laryngeal abduction/adduction is shown in the transillumination records. Evidence of inappropriate glottal abduction/adduction gestures is noted preceding each test word, and also between words in the carrier phrase. Figure 9 shows representative samples from a second profoundly hearing-impaired speaker's production of the same test words. Similar inappropriate glottal gestures between words are again observed. Leaving interarticulator timing for a later discussion, this pattern supports the notion of valving problems at the laryngeal level consistent with previous discussions. During pauses between words, these hearing-impaired speakers inappropriately opened the glottis, a pattern never observed in the production of hearing speakers. Whether these hearing-impaired speakers actually took a breath or simply wasted air cannot be directly ascertained from these data since simultaneous monitoring of respiratory activity was not done. However, the authors argue that the latter is more likely since the glottal abduction gesture was smaller and shorter in duration between words than between utterances. This pattern of aberrant laryngeal valving differs from one hypothesized by Stevens, Nickerson, and Rollins (in press). Based on a spectrographic study of deaf children's productions, these authors hypothesized that the glottis is closed during pauses between words.

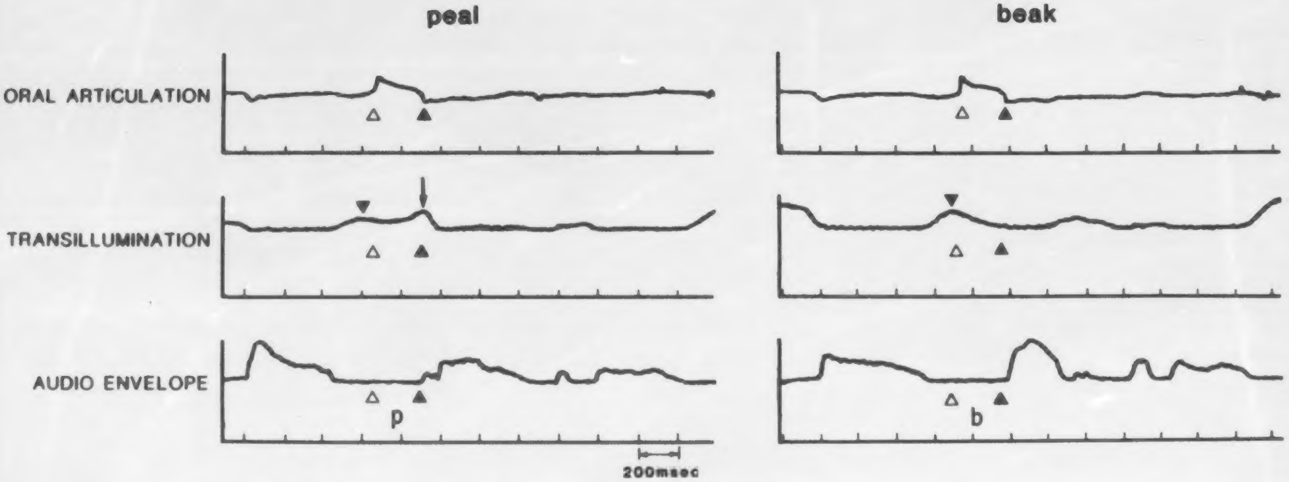


Figure 8. Records of a profoundly hearing-impaired speaker's production of the utterance "Say peal again" (left) and "Say beak again" (right). Curves represent oral articulation (top), transillumination (middle), and audio envelope (bottom). Onset of labial closure for the word initial stops in "peal" and "beak" is marked △, release of oral closure by ▲; ↓ marks peak glottal opening. Examples of inappropriate abduction/adduction gestures are noted in the transillumination record by ▼.

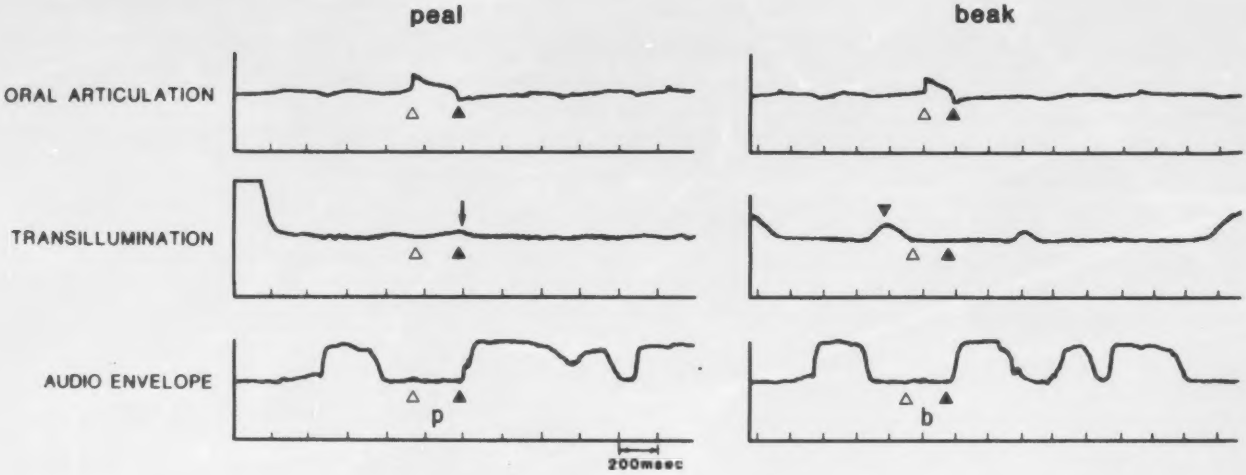


Figure 9. Record of a second profoundly hearing-impaired speaker's productions of the utterances "Say peal again" (left), and "Say beak again" (right). Symbols as in Figure 8.

B. Phonation

The larynx serves as the primary source of acoustic energy for speech and plays an integral role in changes of stress and intonation, and also voicing information. While we have noted earlier in this chapter that hearing-impaired speakers exhibit great difficulty in controlling these phonatory parameters, there are few physiological studies of laryngeal function in the hearing impaired. For convenience of discussion, we will divide laryngeal function into two areas: phonatory and articulatory.

To date, there are few studies that have examined the basic phonatory mechanism in hearing-impaired speakers. One study (Monsen, Engebretson, & Vemula, 1978) examined the glottal volume-velocity waveforms of hearing-impaired speakers using a reflectionless (Sondhi) tube. In this procedure, the subject phonates a neutral vowel into the tube, and a microphone positioned in the tube records a pressure waveform that is considered to be an approximate of the glottal waveform. It should be noted that the use of the Sondhi tube presents some problems in the study of both normal and pathological voice production. In order to provide an accurate estimate of the source waveform, several conditions must be met. For example, the vocal tract itself must have a uniform area function and not contain any side resonators such as the nasal passages. Since inappropriate nasal resonance is a common problem in the speech of the hearing impaired, data obtained using this measurement technique should be interpreted cautiously. Monsen et al. (1978) reported that an individual glottal pulse for a hearing-impaired speaker was not abnormal _per se_, but that differences between hearing-impaired and hearing subjects were seen for successive changes of the glottal waveform from one period to another. Glottal waveforms of hearing-impaired speakers also showed evidence of diplophonia and creaky voice. Thus, the authors hypothesized that hearing-impaired speakers have difficulty controlling overall tension of the vocal folds and also sub-glottal pressure.

Secondly, high-speed laryngeal films have also provided evidence of abnormal laryngeal function in hearing-impaired speakers (Metz, Whitehead, & Mahshie, 1982). Films of several profoundly hearing-impaired speakers show evidence of inappropriate positioning of the vocal folds prior to the onset of phonation and subsequent patterns of abnormal vocal fold vibration. For example, an abnormally high amount of medial compression on the arytenoid cartilages was observed in the films of one hearing-impaired speaker, and only the anterior one-third of the folds vibrated freely. The analysis of these films also revealed that some hearing-impaired speakers do not use appropriate abduction/adduction gestures in producing VCV utterances where C was a voiceless consonant. These data speak to the point of difficulty in laryngeal articulation, that is, the production of segments requiring control and coordination of the larynx.

Laryngeal articulation in the speech of the hearing impaired has been examined in two physiological studies; the first a fiberoptic study of voiced and voiceless segments (Mahshie, 1980) and the second a transillumination study of obstruent production described previously (McGarr & Löfqvist, in press). We have noted in other sections of this chapter that there is considerable evidence from the descriptive and acoustic literature to suggest that hearing-impaired speakers have great difficulty coordinating laryngeal and oral articulatory gestures. One common problem that illustrates this difficulty is confusion of the voiced-voiceless distinction.

Let us consider what is required in the production of a voiceless obstruent (a plosive, fricative or affricative) in the speech of normals. In addition to the supralaryngeal adjustments used to make the closure or constriction, a laryngeal abduction/adduction gesture normally occurs to stop glottal vibration and assist in the build-up of oral pressure. Production of these segments thus involves simultaneous activity at both laryngeal and supralaryngeal levels, and the laryngeal and oral articulations must be coordinated in time. Variations in this timing are used in a wide variety of languages to produce contrasts of voicing and aspiration (cf. Löfqvist & Yoshioka, 1981).

An example of how this interarticulator timing might be manifested in the speech of a hearing subject is shown in Figure 10. Data are taken from the transillumination study of McGarr and Löfqvist described above. At this time, we focus on the temporal relationship between the oral and laryngeal events. Peak glottal opening for the voiceless /p/ in "peal," shown in the transillumination signal (middle record) occurs at about the same time as the end of lip closure (top record) and the burst-release in the acoustic envelope (bottom record). The pattern for this plosive is essentially the same as that obtained for other hearing speakers' production of obstruents in different and unrelated languages (Löfqvist & Yoshioka, 1981). For production of the voiced /b/ in "beak," there is no evidence of glottal opening in the transillumination signal as would be expected for a correct production of this segment.

Figure 11 is an example of a common voiced for voiceless substitution in the speech of a hearing-impaired talker. In this example, the error is due to inappropriate positioning of the vocal folds. For production of the /p/, the transillumination signal shows no evidence of a glottal opening following the onset of lip closure, or any evidence of a burst-release in the acoustic signal. Indeed, listeners judged this production to be a /b/ for /p/ substitution. McGarr and Löfqvist reported that hearing-impaired speakers differed from normal by either omitting the glottal gesture entirely as illustrated above, or by producing a glottal gesture when none was required (see also above, Figures 8 and 9). In fact, one speaker consistently differentiated plosives from fricatives by producing the former without a glottal gesture, but the latter with an opening gesture. However, even when an appropriate laryngeal gesture was made by the hearing-impaired subjects, the timing relative to the oral articulatory events was not normal. Similar observations on the nature of laryngeal articulation have been made by Mahshie (1980). The data from these two studies suggest that hearing-impaired speakers have difficulty in coordinating the temporal and spatial demands of different articulators. We now turn to some evidence that shows that this difficulty in coordination also occurs at the articulatory level.

C. Articulation

Articulatory errors in the speech of the hearing-impaired have been reviewed above. The error patterns described in the literature suggest several hypotheses concerning the physiological "underpinnings" of articula-

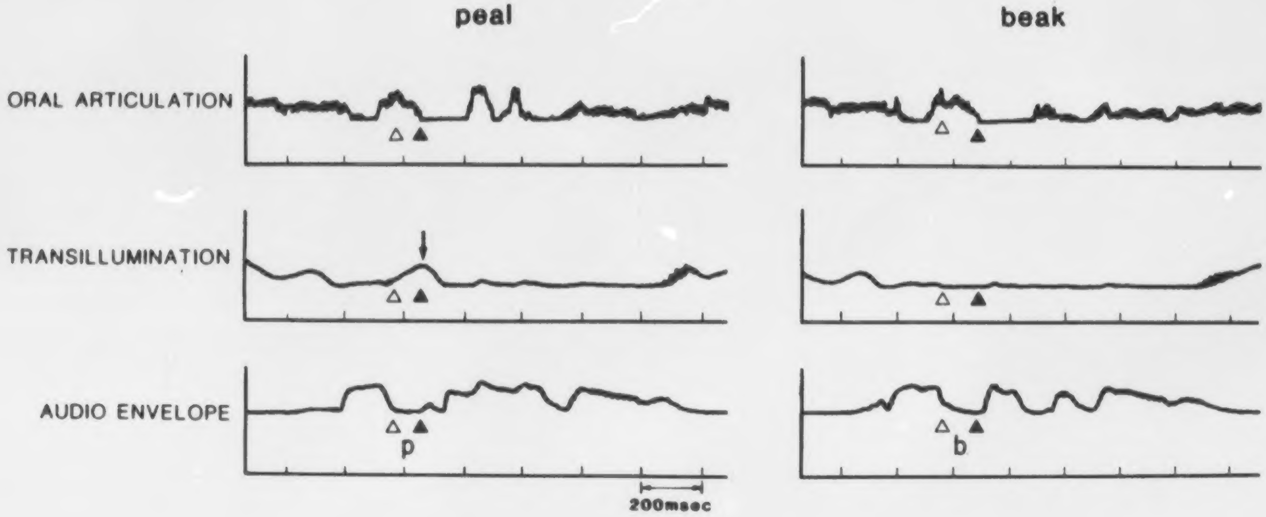


Figure 10. Records of a hearing speaker's productions of the utterances "Say peal again" (left) and "Say beak again" (right). Symbols as in Figure 8. (After McGarr & Löfqvist, in press).

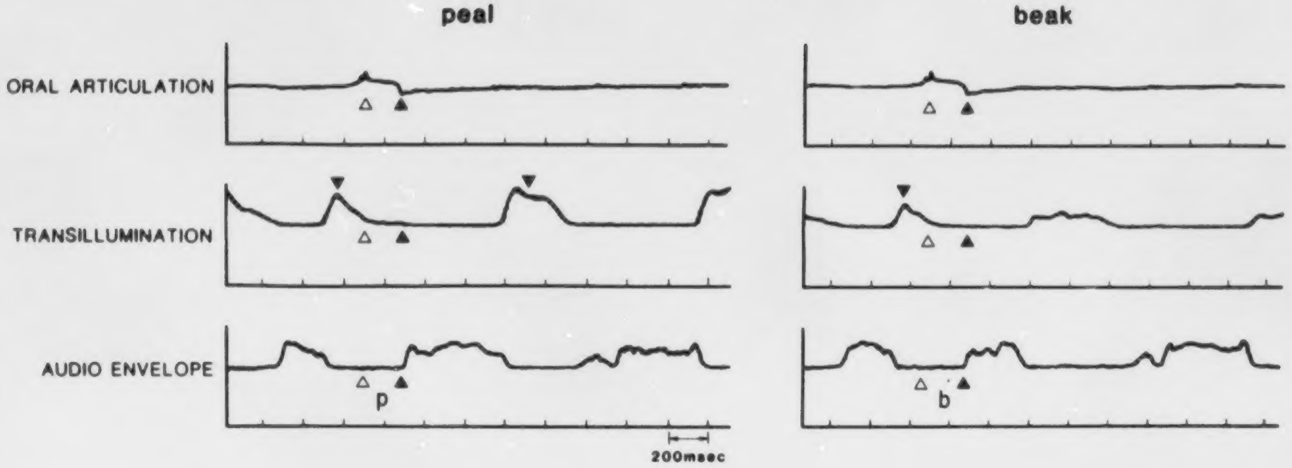


Figure 11. Records of a third profoundly hearing-impaired speaker's production of the utterances "Say peal again" (left), and "Say beak again" (right). Symbols as in Figure 8. This example illustrates the common b/p substitution. (After McGarr & Löfqvist, in press).

tion in the hearing impaired. One hypothesis, derived primarily from studies of consonant production, suggests that hearing-impaired speakers place their articulators fairly accurately but fail to coordinate interarticulator movements. These errors may be broadly characterized as errors in timing. Another hypothesis primarily concerned with vowel articulation is that hearing-impaired speakers move their articulators through a relatively restricted range, thereby "neutralizing" vowels. Again, there have been relatively few physiological studies of articulation in hearing-impaired speakers, three electromyographic investigations (Huntington et al., 1968; McGarr & Harris, 1980; Rothman, 1977), and two cinefluorographic studies (Stein, 1980; Zimmerman & Rettaliata, 1981). These investigations provide some insight into the complex nature of articulatory errors in the hearing impaired.

For example, electromyographic studies of the speech of hearing-impaired persons give ample evidence of instability of production and failure to achieve the tight temporal coupling in articulatory muscles. McGarr and Harris (1980) have shown that for normal speakers, the relationship between two articulators, the lips (orbicularis oris) and the tongue (genioglossus), is closely coordinated in time, and that even changes in stress from one syllable to another do not disrupt this temporal relationship. Indeed, this closely timed interarticulator relationship seems to be characteristic of normal speech production and is evidenced in many articulatory muscles across changes in stress as well as speaking rate (Tuller, Harris, & Kelso, 1981).

Figure 12, taken from electromyographic records of a hearing speaker in the McGarr and Harris experiment, illustrates this temporal relationship. These productions are contrasted in Figure 13 with several examples taken from the records of a hearing-impaired speaker. Clearly, these tokens demonstrate considerable variability on the part of the hearing-impaired speaker in coordinating the activity of the tongue with the lips. Occasionally, tongue activity was timed relatively correctly with respect to lip activity. Most often, the hearing-impaired speaker initiated this tongue activity either too early or too late relative to the lips. These samples suggest that the hearing-impaired speaker does not produce a "wrong" pattern in a stereotypic way; rather, productions are variable from token to token not only for utterances perceived as correct, but also for utterances perceived as incorrect. It is interesting that this variability in production is observed primarily in the lingual rather than the labial component, that is, it is the less visible aspect of articulation that varies. Similar observations have been made regarding phoneme visibility in earlier EMG studies (Huntington et al., 1968; Rothman, 1977). However, observations on the variability in both perceptually correct and incorrect productions clearly provide new insights into the organization of the speech of hearing-impaired talkers.

Cinefluorographic studies (Stein, 1980; Zimmerman & Rettaliata, 1981) provide additional information on upper articulatory movements in hearing-impaired speakers. These X-ray films have been analyzed for an adventitiously hearing-impaired speaker in the former study, and also for five pre-lingually hearing-impaired adults in the latter work. Despite differences in onset of hearing loss, these subjects showed patterns of articulatory dynamics similar to each other, and not unlike normals in many respects. This is not surprising since all of the hearing-impaired speakers were at least partially intelligible. Some of the differences between normal and hearing-impaired

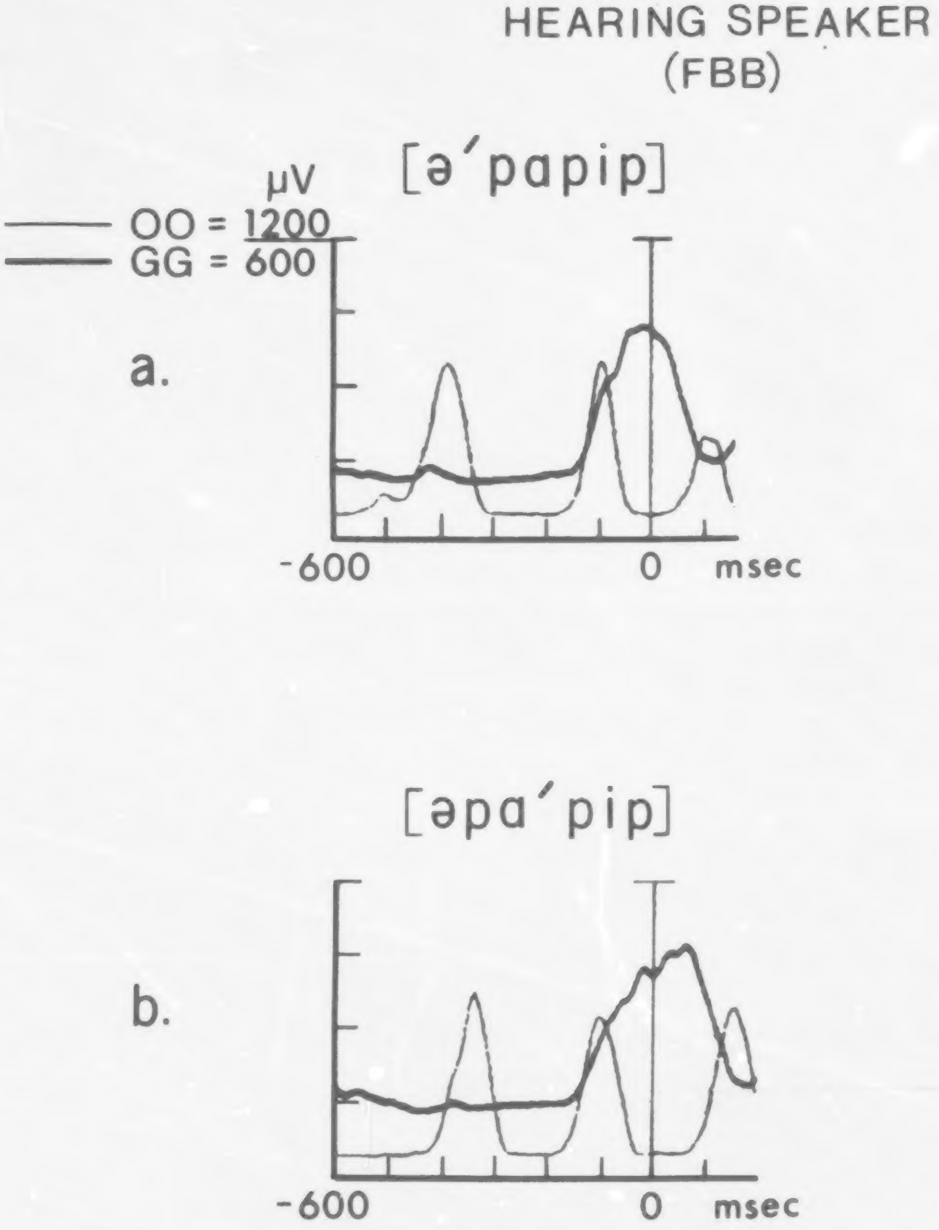


Figure 12. Ensemble average of the EMG potentials for genioglossus (GG) and orbicularis oris (OO) for the utterance (əpapip) produced by a hearing speaker. Stress occurs on V_1 in 5a, or V_2 in 5b, respectively. The vertical line indicates the acoustic release of the /p/ closure. Peak genioglossus activity for the vowel occurs at about the same time as the acoustic burst release (after McGarr & Harris, 1980).

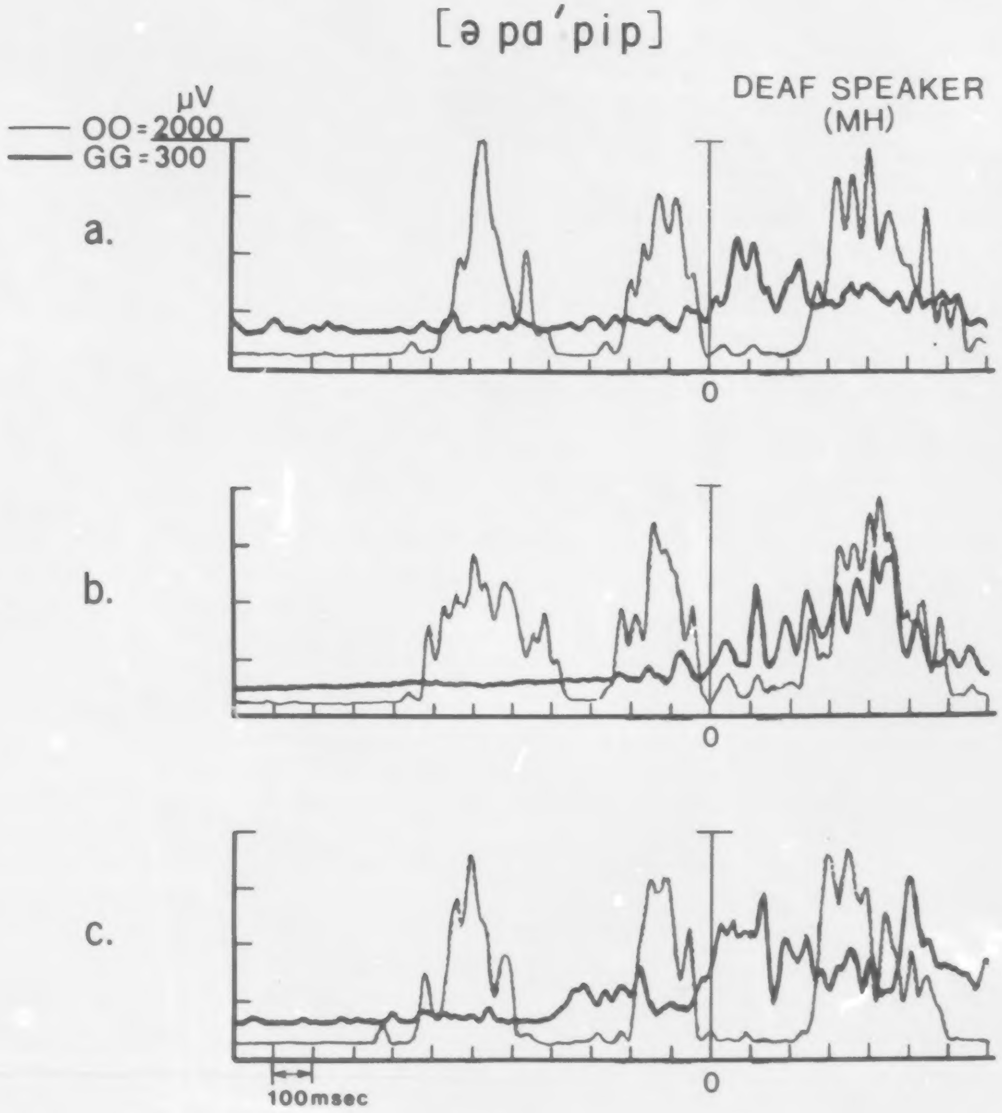


Figure 13. Three selected examples of the EMG potential for the genioglossus and orbicularis oris for the utterance /əpapip/ produced by a profoundly hearing-impaired speaker. The vertical line indicates the acoustic release of the /p/ closure. In Figure 13a, peak genioglossus activity occurs between the second and third orbicularis oris peaks, but it is late relative to the acoustic event. This pattern was most like normal. In Figures 13b and 13c, the single tokens show that genioglossus activity was either too late or too early, respectively (after McGarr & Harris, 1980).

speakers were as follows. Hearing-impaired speakers frequently exhibited faster articulatory speeds for lip, tongue, and jaw movements, and articulatory displacements were often of greater magnitude than for normals. Vowel height differentiation was achieved primarily by jaw movements, and deviant positioning of the tongue, with primarily "fronting" of back vowels noted. A consistent finding of these studies was that onset and offset voicing in consonant production was frequently too long. These physiological data agree with descriptive studies on voicing errors, particularly that of Millin (1971). These results reaffirm the notion that interarticulator coordination is poorly controlled by hearing-impaired speakers.

D. Summary

Taken together, these studies suggest that the physiological correlates of segmental and suprasegmental errors in the speech of the hearing impaired are exceedingly complex. Our knowledge of the physiology of speech production by the hearing impaired is still in the germinal stages. While the research described above has included only relatively few hearing-impaired speakers, and caution must be taken not to overgeneralize results, several interesting mechanisms of production are beginning to emerge.

One is that certain physiological characteristics of the production of hearing-impaired speech may span an entire utterance, and thus cannot be accurately ascribed to either segmental or suprasegmental attributes of speech. These have been termed postural characteristics by Stevens and his colleagues (Stevens, Nickerson, & Rollins, 1978, in press). Examples of postural errors would include inappropriate respiratory control, glottal abduction/adduction gestures, vocal fold tension and mass, tongue position and range of movements, velopharyngeal posture and movements. These postural characteristics include not only the preparatory state for speaking, but also the configuration of the speech production mechanism over time. We have noted several examples in the preceding discussion that suggest that hearing-impaired speakers evidence such inappropriate postures.

The importance of postural characteristics has also been highlighted recently in studies of speech production in normals. Parallels between coordinated non-speech and speech activities have been drawn. For example, a non-speech activity such as locomotion is said to be like speech in that both may be thought of as having a series of rapid, rhythmic, and highly coordinate movements superimposed on a broad posture base. We might think then of speech as a complex and rapidly changing articulatory-phonatory process overlayed on a slowly changing respiratory base. Thus, the hearing-impaired speaker who adopts an inappropriate respiratory posture for whatever reason may preclude the coordination and control of movement elsewhere in the speech production mechanism. An inappropriate respiratory posture may be further exacerbated by inappropriate glottal gestures, or inappropriate tongue position, and so on.

A second problem evidenced by many hearing-impaired speakers is great difficulty in coordinating respiration, phonation, and articulation. In normal speech production, the tight temporal coordination of these events constitutes an important component in any theory of speech production. In the speech production of the hearing impaired, we have ample evidence for a breakdown in interarticulator coordination, for example, in the studies of aerodynamics, laryngeal-supralaryngeal coordination, and articulation cited above. These data suggest not only difficulty accommodating the demands of speech in space and time, but also substantial variation in production from utterance to utterance. Without such coordination, intelligible speech is impossible and taken together, these factors suggest some reasons why listeners find the speech of the hearing impaired so difficult to understand.

Neither problems of postural characteristics nor those of interarticulatory coordination are mutually exclusive. Physiological research focusing on several levels of speech production may prove fruitful in clarifying many of the errors documented in the descriptive literature. A better understanding of these problems at the physiological level will hopefully lead to the development of more effective assessment techniques and training programs for hearing-impaired speakers.

VII. SPEECH INTELLIGIBILITY

We shall use the term "speech intelligibility" to refer to how much of what a child says can be understood by a listener. On the average, the intelligibility of profoundly hearing-impaired children's speech is very poor. Only about one in every five words they say can be understood by a listener who is unfamiliar with the speech of this group (Brannon, 1966; John & Howarth, 1965; Markides, 1970; McGarr, 1978; Smith, 1975).

Before we proceed to a discussion of factors that have been found to affect intelligiblity, some comments on analysis techniques are necessary. First, intelligibility measures in most studies have been based only on a listener's auditory judgments of a child's productions. While this approach may be the most appropriate for quantifying the intelligibility of speech, it does not necessarily provide an accurate assessment of a child's ability to communicate in a face-to-face situation.

A second point that should be made is that the majority of investigators who have attempted to determine the effect of specific variables on intelligibility have done so using a correlational analysis, a statistical analysis of the association between the factor of interest and the reduction in intelligibility. Correlations should be interpreted carefully because a cause and effect relationship cannot be inferred from the results. Several studies that have been performed will be presented in some detail in this section.

A. Hearing Level

A review of the literature indicates that an important factor in determining the intelligibility of a hearing-impaired child's speech is the degree of the child's hearing loss (Boothroyd, 1969; Elliot, 1967; Markides, 1970; Montgomery, 1967; Smith, 1975). Boothroyd (1969) found a correlation between percent intelligibility scores and hearing level at all frequencies, particularly at 1000 Hz and 2000 Hz, for a population of hearing-impaired children from the Clarke School for the Deaf. In fact, the data formed a bimodal distribution: the children with good speech intelligibility (intelligibility score of 70% or more) had considerable hearing, while those children with poor intelligibility (70% or less) had little residual hearing. The median hearing level of the group with good speech intelligibility was 90 dB and, as the hearing loss exceeded 90 dB at 1000 Hz, the median speech scores fell rapidly. In another study that analyzed the speech intelligibility of profoundly hearing-impaired children, Smith (1975) observed a systematic decrease in intelligibility with poorer hearing level to a level of about 85 dB HTL, after which the relationship was not clear. Monsen (1978) found that all the children he studied with hearing losses of 95 dB HTL or less had intelligible speech, but children with losses greater than 95 dB HTL did not always have poor or unintelligible speech. These data indicate that even though a child has a profound hearing loss, he or she still has the potential to develop functional speech skills.

Two studies of interest are those by Smith (1975) and Gold (1978), which were described in the preceding section. Recall that the same test materials and procedures were used in the two studies to assess the speech of different populations of hearing-impaired children. The average intelligibility of the profoundly hearing-impaired children's speech in an oral day school for the deaf was reported by Smith to be about 19%. Gold (1978) reported an average intelligibility score of 39% for the mainstreamed profoundly hearing-impaired children assessed in her study. Thus, children with similar hearing levels in different educational settings showed an average difference of 20% in their intelligibility scores. Not unexpectedly, research has shown that the intelligibility of hard-of-hearing children's speech is substantially higher than that of profoundly hearing-impaired children. Average intelligibility scores of 70-76% have been reported for the hard of hearing (Gold, 1978; Markides, 1970).

Higher intelligibility scores than those mentioned above have been reported by Monsen (1978). His results revealed an average intelligibility score of 91% for severely hearing-impaired children, and a score of 76% for the profoundly hearing-impaired children in his study. Monsen (1978) has attributed the difference in intelligibility scores between his and other studies to differences in the speech material that the children were required to produce. According to Monsen (1978), the sentences in his study were shorter, contained a more familiar vocabulary, and were syntactically less complex than those used by other investigators. In fact, McGarr (1980) has shown that intelligibility scores for hearing-impaired speakers may vary considerably depending on speech material (sentences or words), amount of context, phonetic composition, and, of course, experience of the listener.

The above studies indicate that although the degree of hearing loss is an important variable, this measure alone cannot reliably predict the intelligibility of a child's speech. In fact, in a study by Smith (1975), hearing level was found to be only a fair predictor of the speech intelligibility of profoundly hearing-impaired children. The hearing measure found to be most closely correlated with speech intelligibility was performance on an auditory phoneme recognition test. This finding suggests that it is not hearing level *per se* that is most important for the development of intelligible speech, but rather the ability of the hearing-impaired child to make use of the acoustic cues that are available to him or her.

B. Segmental Errors

It has generally been found that as the overall frequency of segmental or phonemic errors increases in the speech of the hearing impaired, intelligibility decreases (Brannon, 1966; Gold, 1978; Hudgins & Numbers, 1942; Markides, 1970; Smith, 1975). However, the number of segmental errors alone cannot account for reduced intelligibility. Smith (1975), for example, observed that some of the subjects in her study who had approximately the same frequency of segmental errors had speech intelligibility scores differing by as much as 30 percent. Smith hypothesized that these differences appeared to be related, in part, to certain suprasegmental errors that interacted in a complex manner with the segmental errors to reduce intelligibility.

The relationship between specific types of segmental errors and intelligibility has been examined to some extent by Hudgins and Numbers (1942) and later by Smith (1975). In their classic study, Hudgins and Numbers found a high negative correlation between intelligibility and total number of vowel errors (-.61) and total number of consonant errors (-.70). Similar results were reported by Smith, except that she found a slightly higher correlation between vowel errors and intelligibility than did Hudgins and Numbers.

Of the seven consonant error categories considered in the Hudgins and Numbers (1942) study, three categories (omission of initial consonants, voiced-voiceless confusions, and errors involving compound consonants) had the most significant effect on intelligibility. The other four categories considered (substitution errors, nasality errors, omission of final consonants, and errors involving abutting consonants) had a lower correlation with intelligibility and contributed to a much lesser extent to the reduced intelligibility of hearing-impaired children's speech.

In a recent study, Monsen (1978) examined the relationship between intelligibility and four acoustically measured variables of consonant production, three acoustic variables of vowel production, and two measures of prosody. A multiple regression analysis showed that three variables had a high multiple correlation (.85) with intelligibility and thus accounted for 73% of the variance: (1) the difference in voice-onset time between /t/ and /d/; (2) the difference in second formant location between /i/ and /I/, and (3) acoustic characteristics of the nasal and liquid consonants. The first two variables accounted for almost 69% of the variance.

Other segmental errors that have been observed to have a significant negative correlation with intelligibility are: omission of phonemes in the word-initial and medial position; consonant substitutions involving a change in the manner of articulation; substitutions of non-English phonemes such as the glottal stop, and unidentifiable or gross distortions of the intended phoneme (Levitt et al., 1980).

C. Suprasegmental Errors

The suprasegmental errors examined most extensively in relation to intelligibility have been those involving timing. One of the earliest attempts to determine the relationship between deviant timing patterns and intelligibility is found in the study by Hudgins and Numbers (1942). Although they correlated rhythm errors with intelligibility, many of these errors appear to be due to poor timing control. They found that sentences spoken with correct rhythm were substantially more intelligible than those that were not. The correlation between speech rhythm and intelligibility was .73, which was similar to the correlation between total consonant errors and intelligibility, and higher than the correlation found between vowel errors and intelligibility.

The results of other correlational studies have typically shown a moderate negative correlation between excessive prolongation of speech segments and intelligibility (Monsen & Leiter, 1975; Parkhurst & Levitt, 1978). In a recent study, Reilly (1979) found that relative duration (stressed:unstressed syllable nuclei duration ratio) demonstrated a systematic relationship with intelligibility. Reilly (1979) suggested that the better able the profoundly hearing-impaired speaker was to produce the segmental, lexical, and syntactic structure of the utterance, the more intelligible the utterance was likely to be. Data reported by Parkhurst and Levitt (1978) indicate that another type of timing error, the insertion of short pauses at syntactically appropriate boundaries, had a positive effect on intelligibility; the presence of these pauses actually helped to improve intelligibility.

Studies that have attempted to determine the cause and effect relationship between speech errors and intelligibility have dealt primarily with timing. These causal studies can be sub-divided into two major categories: studies in which hearing-impaired children receive intensive training for the correction of timing errors, and studies in which timing errors are corrected in hearing-impaired children's recorded speech samples using modern signal-processing techniques.

The classic training study that attempted to determine the causal relationship between timing errors and intelligibility was conducted by John and Howarth (1965). These investigators reported a significant improvement in the intelligibility of profoundly hearing-impaired children's speech after the children had received intensive training focusing only on the correction of timing errors. In contrast to this, Houde (Note 2) observed a decrement in intelligibility when timing errors of hearing-impaired speakers were corrected, and the results of a similar study by Boothroyd et al. (1974) were equivocal.

A major problem with the training studies is that the training may result in changes in the child's speech other than those of interest. Recent investigations have attempted to eliminate this confounding variable by using computer processing techniques. In these studies, speech is either synthesized with timing distortions, or synthesized versions of the speech of the hearing impaired are modified so that timing errors are corrected. Lang (1975) used an analysis-synthesis approach to correct timing errors in the speech samples produced by hearing-impaired speakers, and also to introduce timing distortions in the samples of normal speakers. Minimal improvements in intelligibility were observed for the speech of the hearing impaired, and minimal decrements in intelligibility were observed for the normal speakers. Bernstein (1977), however, found no reduction in the intelligibility of speech samples produced by a normal speaker when synthesized with timing errors. On the other hand, Huggins (1978) found that when normal speech was synthesized with the durational relationship between stressed and unstressed syllables reversed, there was a substantial reduction in intelligibility. Even greater reductions in intelligibility occurred when the stress assignments for both pitch and duration were incorrect.

In an attempt to resolve some of the conflicting information in this area, Osberger and Levitt (1979) quantified the relative effect of timing errors on intelligibility by means of computer simulation. Speech samples produced by hearing-impaired children were modified to correct timing errors only, leaving all other aspects of the speech unchanged. Three types of correction were performed: relative timing, absolute syllable duration, and pauses. Each error was corrected alone and together with one of the other timing errors. An average improvement in intelligibility was observed only when relative timing errors alone were corrected. The improvement, however, was very small (4%). Since the timing modifications for this condition involved only the correction of the duration ratio for stressed-to-unstressed vowels, the overall durations of the vowels (and syllables) were still longer than the corresponding durations in normal speech. These data indicate that the prolongation of syllables and vowels, which is one of the most obvious deviancies of the speech of the hearing impaired, does not in itself have a detrimental effect on intelligibility.

Attempts have also been made to determine the relationship between errors involving fundamental frequency (Fo) control and intelligibility. Monsen (1978) found that there was no clear-cut relationship between mean Fo and mean amount of Fo change and intelligibility. In their study, McGarr and Osberger (1978) found that, for the majority of the children studied, there seemed to be no simple relationship between pitch deviancy and intelligibility. Some children whose pitch was judged appropriate for their age and sex had intelligible speech, while others did not. The exception to this pattern were the children who were unable to sustain phonation and whose speech contained numerous pitch breaks. Their speech was consistently judged to be unintelligible. Smith (1975) also found that errors involving poor phonatory control (intermittent phonation, spasmodic variations of pitch and loudness, and excessive variability of intonation) had a high correlation with intelligibility.

Data obtained by Parkhurst and Levitt (1978) also suggest that excessive variations in pitch may reduce intelligibility. In this study, a multiple linear regression analysis was performed, relating intelligibility to various prosodic distortions judged to occur in the speech of hearing-impaired children. Breaks in pitch were one of the prosodic errors showing a significant negative regression with intelligibility. The effect of the less deviant patterns, such as elevated Fo, has not been clearly established, although preliminary data suggest that these problems will not have a serious effect on intelligibility.

In summary, we have relatively little information regarding the effect of errors, or combination of errors, on the intelligibility of hearing-impaired children's speech, nor are we able to predict reliably if a child has the potential to develop intelligible speech. Some background variables appear to be important, such as the hearing status of the parent, while others, such as age of identification of hearing impairment, hearing aid use, start of special education, IQ, and the hearing status of siblings show little or no correlation with speech intelligibility (Smith, 1975).

VII. CONCLUDING COMMENTS

We shall now summarize some of the major points discussed in the chapter, and discuss the implications of the available data for the development of assessment and training techniques. On the basis of the data presented, the following statements can be made regarding the speech production skills of hearing-impaired children:

1. Rate of vocal output cannot be used to describe accurately the differences in the vocalization behavior between hearing and hearing-impaired infants. Striking differences between the vocalizations of normal-hearing and hearing-impaired infants do emerge at an early age, but the differences are seen in phonemic production rather than rate of vocal output. Specifically, hearing-impaired infants tend to produce stereotypic vocalization patterns with a reduced phonemic repertoire relative to hearing infants.

2. The developmental stages of speech acquisition in the hearing impaired appear to be similar to those of normal-hearing children in some respects but not in others. Also, the speech production patterns of older hearing-impaired children show many similarities to the patterns of the younger hearing-impaired children.

3. Segmental errors, as determined by phonetic transcriptions of hearing-impaired children's speech can be classified by the following two categories:

   a) Omission Errors: This type of error most often involves consonants, particularly those in the word-final position. Omission of vowels is infrequent and usually does not occur unless the entire syllable has been omitted.

   b) Substitution Errors: Frequent errors in this category involve confusion between voiced-voiceless cognates, substitution of a consonant with the same place of production but a different manner of production as the intended consonant (and vice versa), and substitution of non-English sounds, particularly the glottal stop for the intended phoneme. Vowel errors in this category typically involve tense-lax substitutions, substitutions toward a vowel that is more central than the target vowel, and substitution of the schwa vowel for the intended vowel. Diphthong errors frequently involve substitutions of one of the elements of a closely related vowel.

4. Errors are less frequent for consonant phonemes produced at the front of the mouth (the labial and labio-dental consonants) as compared to phonemes with a place of articulation at the middle or back of the mouth. Traditionally, this pattern of production has been attributed to the greater visibility of phonemes produced in the front of the mouth. Other articulatory considerations, such as the relatively constrained movements of the most visible articulators, the lips, may also account for this production pattern.

5. Similar error patterns have been found to occur in the speech of different groups of hearing-impaired children. The largest difference between children is in the frequency of errors; type of error may also vary, but to a lesser extent than frequency of errors.

6. At the suprasegmental level of production, poor timing control produces the following deviations:

a) Prolongation of speech segments

b) Distortion of temporal relationship between speech segments

c) Insertion of frequent and lengthy pauses often at syntactically inappropriate boundaries

d) Distortion of phonetic context effects

e) Insertion of adventitious phonemes.

Poor control of fundamental frequency can result in problems such as:

a) Average pitch level too high

b) Intonation with insufficient variability

c) Intonation with excessive variability.

Abnormal voice characteristics such as harshness, breathiness, hyper- and hyponasality may also be present.

7. Acoustic analyses have shown manifestations of the above perceptual errors in the distortion of voice-onset-time, formant frequency transitions, frequency location of the formants, and segmental durations.

8. Recent studies have begun to detail the physiological correlates of segmental and suprasegmental errors. These studies show that the underlying causes of error patterns are more complex than has been alluded to in the descriptive literature. Some of the production mechanisms responsible for the perceptual and acoustic distortions are poor respiratory control, evidenced by initiation of phonation at too low a level of vital capacity and production of a reduced number of syllables per breath; abnormal laryngeal function, evidenced by laryngeal valving problems and failure to coordinate laryngeal and respiratory events, and a breakdown in interarticulator programming, evidenced by poor control and coordination of articulatory gestures, both at the laryngeal and supralaryngeal levels of production. Improper postural characteristics of the speech mechanism may affect many aspects of speech production and result in segmental and suprasegmental misperceptions.

9. Although there are many deviations in the speech of the hearing-impaired, these deviations do not generally occur in a random way. There is evidence that many of the deviations are phonetically and phonologically consistent albeit the systems may not be the same as those used by normal-hearing talkers. However, the use of a deviant phonological system will still

pose problems for the listener who must decode the intended message. Data are also available suggesting that hearing-impaired talkers manipulate some segmental, lexical, and syntactic aspects of speech in the same manner as normals.

10. The intelligibility of the speech of children with profound hearing losses in day schools for the hearing impaired has been reported to be about 20%. This figure is based on the percentage of words correctly understood through audition alone by persons who are unfamiliar with the speech of the hearing impaired. Under the same conditions, the intelligibility of the speech of children with profound losses who are mainstreamed has been found to be about 40%. The intelligibility of the speech of hard-of-hearing children is substantially higher than that of severely and profoundly hearing-impaired children.

11. The intelligibility of hearing-impaired children's speech has been found to be influenced by the degree of linguistic context and the experience of the listener with the speech of the hearing-impaired.

12. The relationship between specific error types and intelligibility has not been clearly established. Correlational studies show a high degree of association between the frequency of segmental errors and reduction in intelligibility. Of the various error types that have been studied, the highest correlations have been reported for overall frequency of phonemic errors, errors of omission in the word-initial and medial position, substitutions involving a change in the manner of articulation. substitution of non-English phonemes, and unidentifiable or other gross distortions of the intended phonemes. At the suprasegmental level, timing errors and errors involving poor phonatory control have been found to have a negative effect on intelligibility.

Although our knowledge about the speech of the hearing impaired is far from complete, implications for assessment and training strategies can be gleaned from the aforementioned findings. First, hearing-impaired children show distinct error patterns, and unless appropriate assessment instruments are used, some errors may go undetected. Second, in addition to assessing speech structures, clinicians and teachers must attempt to evaluate the adequacy of respiratory, laryngeal, and articulatory maneuvers essential for normal speech production. By this, we do not mean to imply that physiological measures should be performed routinely in the clinic. Rather, through clinical observation and perceptual measures, inferences can be made about the underlying speech production mechanism. Third, a phonological analysis of an individual child's sound system will enable the clinician to determine if a child's speech deviates from normal in a systematic way, or if the errors are random.

Following the evaluation, the clinician or teacher should raise pertinent questions regarding each child's error patterns and production skills. Such questions include the following:

1. Does the child have a diverse sound system?

    a) Are the basic contrasts, i.e., oral-nasal, stop-continuant, fricative-nonfricative, present in the child's sound system?

    b) Are these contrasts present for the different places of articulation, i.e., front, mid, back?

    c) Is there vowel differentiation, i.e., front-back contrast, high-low contrast?

    d) Are non-English sounds (glottal stop) or unidentifiable sounds frequently substituted for the intended phoneme?

2. Is there adequate control of the speech mechansim?

    a) Is there adequate breath management? Is the feature of frication absent or distorted; is there evidence of phrase structure, with or without a terminal fall in pitch?

    b) Is there poor velopharyngeal control that results in segmental errors (substitution of oral sounds for nasal sounds), and an abnormal voice quality (hypernasality)?

    c) Is there adequate laryngeal control? Are there excessive changes in pitch, are there inappropriate changes in pitch? Are there localized changes in fundamental frequency that are not linked appropriately to changes in lexical stress?

    d) Is there coordination between laryngeal and supralaryngeal movements, i.e., are there voiced-voiceless errors?

    e) Is there independent control of vowel production and pitch control? Is there a noticeable difference in pitch between productions of low vowels /æ, a/ and high vowels /i, u/?

    f) Is there adequate timing control? Is overall rate too slow; are there adventitious sounds; are there distortions of temporal relationships between segments and distortion of phonetic context effects in the temporal domain; are pauses frequently inserted; is there glottalization?

Once these areas are addressed, an optimal training sequence can be selected to meet the individual needs of each child. The effectiveness of the training strategies can be assessed through careful and objective monitoring of the child's performance in speech therapy.

REFERENCE NOTES

1. Honda, K. *Relationship between pitch control and vowel articulation.* Paper presented at the "Vocal Fold Physiology Conference," University of Wisconsin, 1981.

2. Houde, R. A. *Instantaneous visual feedback in speech training for the deaf*. Paper presented at the American Speech and Hearing Association Convention, 1973.

REFERENCES

Angelocci, A., Kopp, G., & Holbrook, A. The vowel formants of deaf and normal hearing eleven to fourteen year old boys. *Journal of Speech and Hearing Disorders*, 1964, *29*, 156-170.

Bernstein, J. Intelligibility and simulated deaf-like speech. *Conference Record*, IEEE International Conference on Acoustics, Speech, and Signal Processing, Hartford, CT, 1977.

Boone, D. R. Modification of the voices of deaf children. *The Volta Review*, 1966, *68*, 686-692.

Boothroyd, A. Distribution of hearing levels in the student population of the Clarke School for the Deaf. Northampton, Mass.: Clarke School for the Deaf, 1969.

Boothroyd, A., Nickerson, R., & Stevens, K. Temporal patterns in the speech of the deaf: A study in remedial training. Northampton, Mass.: Clarke School for the Deaf, 1974.

Borden, G. J, & Harris, K. S. *Speech science primer: Physiology, acoustics, and perception of speech*. Baltimore: Williams & Wilkins, 1980.

Brannon, J. B. The speech production and spoken language of the deaf. *Language and Speech*, 1966, *9*, 127-136.

Brehm, F. E. Speech correction. *American Annals of the Deaf*, 1922, *67*, 361-370.

Bush, M. *Vowel articulation and laryngeal control in the speech of the deaf*. Unpublished doctoral dissertation, Massachusetts Institute of Technology, 1981.

Calvert, D. R. *Some acoustic characteristics of the speech of profoundly deaf individuals*. Unpublished doctoral dissertation, Stanford University, 1961.

Carr, J. An investigation of the spontaneous speech sounds of five-year-old deaf-born children. *Journal of Speech and Hearing Disorders*, 1953, *18*, 22-29.

Davis, J. Personnel and services. In J. Davis (Ed.), *Our forgotten children: Hard-of-hearing pupils in the schools*. Minneapolis, Minn.: National Support Systems, 1977, 27-42.

Delattre, P. C., Liberman, A. M., & Cooper, F. S. Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America*, 1955, *27*, 769-773.

Denes, P. Effect of duration on the perception of voicing. *Journal of the Acoustical Society of America*, 1955, *27*, 761-764.

Elliot, L. L. Descriptive analysis of audiometric and psychometric scores of students at a school for the deaf. *Journal of Speech and Hearing Research*, 1967, *10*, 21-40.

Fairbanks, G. Recent experimental investigations of vocal pitch in speech. *Journal of the Acoustical Society of America*, 1940, *11*, 457-466.

Fairbanks, G., Herbert, E. L., & Hammond, J. M. An acoustical study of vocal pitch in seven and eight year old girls. *Child Development*, 1949, *20*, 71-78.

Fairbanks, G., Wiley, J. H., & Lassman, F. M. An acoustical study of vocal pitch in seven and eight year old boys. *Child Development*, 1949, *20*, 63-69.

Forner, L., & Hixon, T. Respiratory kinematics in profoundly hearing-impaired speakers. *Journal of Speech and Hearing Research*, 1977, *20*, 373-408.

Froeschels, E. *Psychological elements in speech*. Boston: Expression Co., 1932.

Fry, D. B. Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America*, 1955, *27*, 765-768.

Fry, D. B. Experiments in the perception of stress. *Language and Speech*, 1958, *1*, 126-152.

Geffner, D. Feature characteristics of spontaneous speech production in young deaf children. *Journal of Communication Disorders*, 1980, *13*, 443-454.

Geffner, D., & Freeman, L. Speech assessment at the primary level: Interpretation relative to speech training. In J. Subtelny (Ed.), *Speech assessment and speech improvement for the hearing impaired*. Washington, D.C.: A. G. Bell Association, 1980, 1-17.

Gold, T. *Speech and hearing skills: A comparison between hard-of-hearing and deaf children*. Unpublished doctoral dissertation, City University of New York, 1978.

Green, D. S. *Fundamental frequency of the speech of profoundly deaf individuals*. Unpublished doctoral dissertation, Purdue University, 1956.

Hasek, C., Singh, S., & Murry, T. Acoustic attributes of pre-adolescent voices. *Journal of the Acoustical Society of America*, 1981, *68*, 1262-1265.

Haycock, G. S. *The teaching of speech*. Stoke-on-Trent, Eng.: Hill and Ainsworth, Ltd., 1933.

Heider, F., Heider, G., & Sykes, J. A study of the spontaneous vocalizations of fourteen deaf children. *The Volta Review*, 1941, *43*, 10-14.

Heidinger, V. A. *An exploratory study of procedures for improving temporal features in the speech of deaf children*. Unpublished Ed.D. Thesis, Teachers College, Columbia University, 1972.

Hixon, T., Goldman, M., & Mead, J. Kinematics of the chest wall during speech production: Volume displacements of the rib cage, abdomen, and lung. *Journal of Speech and Hearing Research*, 1973, *16*, 78-115.

Hixon, T., Mead, J., & Goldman, M. Dynamics of the chest wall during speech production: Function of the thorax, rib cage, diaphragm, and abdomen. *Journal of Speech and Hearing Research*, 1976, *19*, 297-356.

Hollien, H., & Paul, P. A second evaluation of the speaking fundamental frequency and chronological age in males. *Language and Speech*, 1969, *12*, 119-124.

Hood, R. B. *Some physical concomitants of the perception of speech rhythm of the deaf*. Unpublished doctoral dissertation, Stanford University, 1966.

House, A. S. On vowel duration in English. *Journal of the Acoustical Society of America*, 1961, *33*, 1174-1178.

House, A. S., & Fairbanks, G. The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America*, 1953, *25*, 105-113.

Hudgins, C. V. A comparative study of the speech coordinations of deaf and normal-hearing subjects. *Journal of Genetic Psychology*, 1934, *44*, 3-46.

Hudgins, C. V. A study of respiration and speech. *The Volta Review*, 1936, *38*, 341-343, 347.

Hudgins, C. V. Voice production and breath control in the speech of the deaf. *American Annals of the Deaf*, 1937, *82*, 338-363.

Hudgins, C. V. Speech breathing and speech intelligibility. *The Volta Review*, 1946, *48*, 642-644.

Hudgins, C. V., & Numbers, F. C. An investigation of the intelligibility of the speech of the deaf. *Genetic Psychology Monographs*, 1942, *25*, 289-392.

Huggins, A. W. F. Speech timing and intelligibility. In J. Requin (Ed.), *Attention and performance VII*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1978, 279-298.

Huggins, A. W. F. Better spectrograms from children's speech. *Journal of Speech and Hearing Research*, 1980, *23*, 19-27.

Huntington, D., Harris, K. S., & Sholes, G. An electromyographic study of consonant articulation in hearing-impaired and normal speakers. *Journal of Speech and Hearing Research*, 1968, *11*, 147-158.

Hutchinson, J. M., & Smith, L. Aerodynamic functioning in consonant production by hearing-impaired adults. *Audiology and Hearing Education*, 1976, *2*, 16-19, 22-25, 34.

Irvin, B., & Wilson, L. The voiced-unvoiced distinction in deaf speech. *American Annals of the Deaf*, 1973, *118*, 43-45.

John, J. E. J., & Howarth, J. N. The effect of time distortions on the intelligibility of deaf children's speech. *Language and Speech*, 1965, *8*, 127-134.

Klatt, D. H. Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics*, 1975, *3*, 129-140.

Lach, R., Ling, D., Ling, A., & Ship, N. Early speech development in deaf infants. *American Annals of the Deaf*, 1970, *115*, 522-526.

Lang, H. G. *A computer-based analysis of the effects of rhythm modification on the intelligibility of the speech of hearing and deaf subjects*. Unpublished masters thesis, Rochester Institute of Technology, 1975.

Levitt, H., Smith, C. R., & Stromberg, H. Acoustic, articulatory, and perceptual characteristics of the speech of deaf children. In G. Fant (Ed.), *Proceedings of the Speech Communication Seminar, Stockholm*. New York: Wiley, 1974, 129-139.

Levitt, H., Stromberg, H., Smith, C., & Gold, T. The structure of segmental errors in the speech of deaf children. *Journal of Communication Disorders*, 1980, *13*, 419-442.

Liberman, A. M., Delattre, P. C., Gerstman, L. C., & Cooper, F. S. Tempo of frequency change as a cue for distinguishing classes of speech sounds. *Journal of Experimental Psychology*, 1956, *52*, 127-137.

Lisker, L., & Abramson, A. A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 1964, *20*, 384-422.

Löfqvist, A., & Yoshioka, H. Interarticulator programming in obstruent production. *Phonetica*, 1981, *38*, 21-34.

Mahshie, J. *Laryngeal behavior in hearing-impaired speakers*. Unpublished doctoral dissertation, Syracuse University, 1980.

Mangan, K. Speech improvement through articulation testing. *American Annals of the Deaf*, 1961, *106*, 391-396.

Markides, A. The speech of deaf and partially hearing children with special reference to factors affecting intelligibility. *British Journal of Disorders of Communication*, 1970, *5*, 126-140.

Martony, J. Studies on speech of the deaf. *Quarterly Progress and Status Report* (Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden), 1966, 16-24.

Martony, J. On the correction of the voice pitch level for severely hard of hearing subjects. *American Annals of the Deaf*, 1968, *113*, 195-202.

Mavilya, M. P. *Spontaneous vocalization and babbling in hearing impaired infants*. In G. Fant (Ed.), *International symposium on speech communication abilities and profound deafness*. Washington, D.C.: A. G. Bell Association, 1968.

McGarr, N. S. *The differences between experienced and inexperienced listeners in understanding the speech of the deaf*. Unpublished doctoral dissertation, The City University of New York, 1978.

McGarr, N. S. Evaluation of speech in intermediate school-aged deaf children. In J. Subtelny (Ed.), *Special assessment and speech improvement for the hearing impaired*. Washington, D.C.: Alexander Graham Bell Association, 1980, 45-66.

McGarr, N. S., & Harris, K. S. Articulatory control in a deaf speaker. *Haskins Laboratories Status Report on Speech Research*, 1980, *SR-63/64*, 309-332.

McGarr, N. S., & Löfqvist, A. Obstruent production in hearing-impaired speakers: Interarticulator timing and acoustics. *Journal of the Acoustical Society of America*, in press.

McGarr, N. S., & Osberger, M. J. Pitch deviancy and the intelligibility of deaf children's speech. *Journal of Communication Disorders*, 1978, *11*, 237-248.

Metz, D., Whitehead, R., & Mahshie, J. Physiological correlates of the speech of the deaf. In D. Sims (Ed.), *Deafness and communication: Assessment and training*. Baltimore, Md.: Williams and Wilkins, 1982.

Millin, J. Therapy for reduction of continuous phonation in the hard-of-hearing population. *Journal of Speech and Hearing Disorders*, 1971, *36*, 496-498.

Monsen, R. B. Durational aspects of vowel production in the speech of deaf children. *Journal of Speech and Hearing Research*, 1974, *17*, 386-398.

Monsen, R. B. Normal and reduced-phonological space: The production of English vowels by deaf adolescents. *Journal of Phonetics*, 1976, *4*, 189-198. (a)

Monsen, R. B. The production of English stop consonants in the speech of deaf children. *Journal of Phonetics*, 1976, *4*, 29-42. (b)

Monsen, R. B. Second formant transitions of selected consonant-vowel combinations in the speech of deaf and normal-hearing children. *Journal of Speech and Hearing Research*, 1976, *19*, 279-289. (c)

Monsen, R. B. A taxonomic study of diphthong production in the speech of deaf children. In S. K. Hirsh, D. H. Eldredge, I. S. Hirsh, & S. R. Silverman (Eds.), *Hearing and Davis: Essays honoring Hallowell Davis*. St. Louis, Mo.: Washington University Press, 1976, 281-290. (d)

Monsen, R. B. Toward measuring how well hearing-impaired children speak. *Journal of Speech and Hearing Research*, 1978, *21*, 197-219.

Monsen, R. B. Acoustic qualities of phonation in young hearing-impaired children. *Journal of Speech and Hearing Research*, 1979, *22*, 270-288.

Monsen, R. B., Engebretson, A. M., & Vemula, N. Some effects of deafness of the generation of voice. *Journal of the Acoustical Society of America*, 1978, *66*, 1680-1690.

Monsen, R. B., & Leiter, E. Comparison of intelligibility with duration and pitch control in the speech of deaf children. *Journal of the Acoustical Society of America*, 1975, *57*, 569. (Abstract)

Montgomery, G. W. G. Analysis of pure tone audiometric responses in relation to speech development in the profoundly deaf. *Journal of the Acoustical Society of America*, 1967, *41*, 53-59.

Nober, H. Articulation of the deaf. *Exceptional Children*, 1967, *33*, 611-621.

Oller, D. K., Jensen, H., & Lafayette, R. The relatedness of phonological processes of a hearing-impaired child. *Journal of Communication Disorders*, 1978, *11*, 96-106.

Oller, D. K., & Kelly, C. A. Phonological substitutions of a hard-of-hearing child. *Journal of Speech and Hearing Disorders*, 1974, *39*, 65-74.

Osberger, M. J. *The effect of timing errors on the intelligibility of deaf children's speech*. Unpublished doctoral dissertation, The City University of New York, 1978.

Osberger, M. J. Fundamental frequency characteristics of the speech of the hearing impaired. *Journal of the Acoustical Society of America*, 1981, *69*, S69. (Abstract)

Osberger, M. J., & Levitt, H. The effect of timing errors on the intelligibility of deaf children's speech. *Journal of the Acoustical Society of America*, 1979, *66*, 1316-1324.

Osberger, M. J., Levitt, H., & Slosberg, R. Acoustic characteristics of correctly produced vowels. *Journal of the Acoustical Society of America*, 1979, *66*, S13. (Abstract)

Parkhurst, B., & Levitt, H. The effect of selected prosodic errors on the intelligibility of deaf speech. *Journal of Communication Disorders*, 1978, *11*, 249-256.

Parmenter, C. E., & Trevino, S. N. Relative durations of stressed to unstressed vowels. *American Speech*, 1936, *10*, 129-136.

Peterson, G. E., & Barney, H. L. Control methods used in a study of vowels. *Journal of the Acoustical Society of America*, 1952, *24*, 175-184.

Peterson, G. E., & Lehiste, I. Duration of syllable nuclei in English. *Journal of the Acoustical Society of America*, 1960, *32*, 693-703.

Pickett, J. M. *The sounds of speech communication*. Baltimore, Md.: University Park Press, 1980.

Raphael, L. J. Preceding vowel duration as a cue to the perception of the voicing characteristics of word-final consonants in American English. *Journal of the Acoustical Society of America*, 1972, *51*, 1296-1303.

Rawlings, C. A comparative study of the movements of the breathing muscles in speech and quiet breathing of deaf and normal subjects. *American Annals of the Deaf*, 1935, *80*, 147-156.

Reilly, A. P. *Syllabic nucleus duration in the speech of hearing and deaf children*. Unpublished doctoral dissertation, The City University of New York, 1979.

Rothman, H. A spectrographic investigation of consonant-vowel transitions in the speech of deaf adults. *Journal of Phonetics*, 1976, *4*, 129-136.

Rothman, H. An electromyographic investigation of articulation and phonation patterns in the speech of deaf adults. *Journal of Phonetics*, 1977, *5*, 369-376.

Russell, G. Normal speech pitch patterns for the deaf. A new approach. *Oralism and Auralism*, 1929, *8*, 1-17.

Scuri, D. Restirazione e fonazione nei sordomute. *Rassegna di Educazione e fonetica biologica*, 1935, *14*, 82-113.

Scripture, E. W. The voices of the deaf. *The Volta Review*, 1913, *15*, 77-80.

Slis, I. H., & Cohen, A. On the complex regulating the voiced-voiceless distinction I, II. *Language and Speech*, 1969, *12*, 80-102, 137-155.

Smith, C. R. Residual hearing and speech production in deaf children. *Journal of Speech and Hearing Research*, 1975, *18*, 795-811.

Stark, R. E. Vocalizations of the pre-school deaf child. *Progress Report*. Baltimore, Md.: The Johns Hopkins School of Medicine, 1967.

Stark, R. E. Phonatory development in young normally-hearing and hearing-impaired children. In I. Hochberg, H. Levitt, & M. J. Osberger (Eds.), *Speech of the hearing impaired: Research, training, and personnel preparation*. Baltimore, Md.: University Park Press, in press.

Stein, D. *A study of articulatory characteristics of deaf talkers*. Unpublished doctoral dissertation, University of Iowa, 1980.

Stevens, K., Nickerson, R., Boothroyd, A., & Rollins, A. Assessment of nasalization in the speech of deaf children. *Journal of Speech and Hearing Research*, 1976, *19*, 393-416.

Stevens, K. N., Nickerson, R. S., & Rollins, A. M. On describing the suprasegmental properties of the speech of deaf children. In D. McPherson & M. Davids (Eds.), *Advances in prosthetic devices for the deaf: A technical workshop*. Rochester, N.Y.: National Technical Institute for the Deaf, 1978, 134-155.

Stevens, K. N., Nickerson, R. S., & Rollins, A. Suprasegmental and posture aspects of speech production and their effect on articulatory skills and intelligibility. In I. Hochberg, H. Levitt, & M. J. Osberger (Eds.), *Speech of the hearing impaired: Research, training, and personnel preparation*. Baltimore, Md.: University Park Press, in press.

Stoel-Gammon, C. The acquisition of segmental phonology by normal-hearing and hearing-impaired children. In I. Hochberg, H. Levitt, & M. J. Osberger (Eds.), *Speech of the hearing impaired: Research, training, and personnel preparation*. Baltimore, Md.: University Park Press, in press.

Story, A. J. Talks to young teachers of speech: The development of voice for speaking. *The Volta Review*, 1917, *19*, 68-70.

Subtelny, J. D. Speech assessment of the deaf adult. *Journal of the Academy of Rehabilitative Audiology*, 1975, *8*, 110-116.

Sykes, J. L. A study of the spontaneous vacalizations of young deaf children. *Psychological Monograph*, 1940, *52*, 104-123.

Tuller, B., Harris, K. S., & Kelso, J. A. S. Articulatory motor events as a function of speaking rate and stress. *Haskins Laboratories Status Report on Speech Research*, 1981, *SR-65*, 33-62.

Voelker, C. H. A preliminary strobophotoscopic study of the speech of the deaf. *American Annals of the Deaf*, 1935, *80*, 243-259.

Voelker, C. H. An experimental study of the comparative rate of utterance of deaf and normal-hearing speakers. *American Annals of the Deaf*, 1938, *83*, 274-284.

West, J., & Weber, J. A phonological analysis of the spontaneous language of a four year old, hard of hearing child. *Journal of Speech and Hearing Disorders*, 1973, *38*, 25-35.

Whitehead, R. Some respiratory and aerodynamic patterns in the speech of the hearing impaired. In I. Hochberg, H. Levitt, & M. J. Osberger (Eds.), *Speech of the hearing impaired: Research, training, and personnel preparation*. Baltimore, Md.: University Park Press, in press.

Whitehead, R., & Barefoot, S. Some aerodynamic characteristics of plosive consonants produced by hearing-impaired speakers. *American Annals of the Deaf*, 1980, *125*, 366-373.

Zimmerman, G., & Rettaliata, P. Articulatory patterns of an adventitiously deaf speaker: Implications for the role of auditory information in speech production. *Journal of Speech and Hearing Research*, 1981, *24*, 169-178.

II. PUBLICATIONS

III. APPENDIX

PUBLICATIONS

Baer, T. Investigation of the phonatory mechanism. In C. L. Ludlow & M. O. Hart (Eds.), *ASHA Reports: Proceedings of the Conference on the Assessment of Vocal Pathology*, 1981, *11*, 38-46.

Bell-Berti, F., & Harris, K. S. Temporal patterns of coarticulation: Lip rounding. *Journal of the Acoustical Society of America*, 1982, *71*, 449-454.

Best, C. T. Developmental neuropsychology: Hemispheric specialization and integration (Conference Abstracts). *International Journal of Neuroscience*, in press.

Best, C. T., Hoffman, H., & Glanville, B. B. Development of infant ear asymmetries for speech and music. *Perception & Psychophysics*, 1982, *31*, 75-85.

Brady, S., Shankweiler, D., & Mann, V. Speech perception and memory coding in relation to reading ability. *Journal of Experimental Child Psychology*, in press.

Crowder, R. G. The role of auditory memory in speech perception and discrimination. In T. Meyers, J. Laver, & J. Anderson (Eds.), *The cognitive representation of speech*. Amsterdam: North-Holland Publishing Co., 1981.

Crowder, R. G. Decay of information in auditory memory. *Journal of Experimental Psychology: Human Learning and Memory*, 1982, *8*, 153-162.

Crowder, R. G. General forgetting theory and the locus of amnesia. In L. Cermak (Ed.), *Human memory and amnesia*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1982.

Fowler, C. Some aspects of language perception by eye. In O. Tzeng & H. Singer (Eds.), *Perception of print*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1981, 171-196.

Fowler, C., Wolford, G., Slade, R., & Tassinary, L. Lexical access with and without awareness. *Journal of Experimental Psychology: General*, 1981, *110*, 341-362.

Harris, K. S. Electromyography as a technique for laryngeal investigation. In C. L. Ludlow & M. O. Hart (Eds.), *ASHA Reports: Proceedings of the Conference on the Assessment of Vocal Pathology*, 1981, *11*, 70-84.

Healy, A. F. Cognitive processes in reading text. *Cognition*, 1981, *10*, 119-126.

Healy, A. F. The effects of visual similarity on proofreading for misspellings. *Memory & Cognition*, 1981, *9*, 453-460.

Healy, A. F., & Repp, B. H. Context independence and phonetic mediation in categorical perception. *Journal of Experimental Psychology: Human Perception and Performance*, 1982, *8*, 68-80.

Katz, L., & Feldman, L. B. The relationship between pronunciation and recognition of printed words in shallow and deep orthographies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, in press.

Katz, R. B., Shankweiler, D., & Liberman, I. Y. Memory for item order and phonetic recoding in the beginning reader. *Journal of Experimental Child Psychology*, 1981, *32*, 474-484.

Kay, B. A., Kostic, A., & Katz, L. *Decomposition and grammatical priming in lexical decision*. Paper presented at the 53rd annual meeting of the Eastern Psychological Association, Baltimore, Md., April 1981.

Kelso, J. A. S. Contrasting perspectives on order and regulation in movement. In J. Long & A. Baddeley (Eds.), *Attention and performance* (IX). Hillsdale, N.J.: Erlbaum, 1981.

Kelso, J. A. S., & Clark, J. E. (Eds.) *The development of movement control and coordination*. Chichester: John Wiley, 1982.

Kelso, J. A. S., Holt, K. G., Rubin, P., & Kugler, P. N. Patterns of human interlimb coordination emerge from the properties of nonlinear limit cycle oscillatory processes: Theory and data. *Journal of Motor Behavior*, 1981.

Kelso, J. A. S., & Saltzman, E. L. Motor control: Which themes do we orchestrate? *The Behavioral and Brain Sciences*, in press.

Kugler, P. N., Kelso, J. A. S., & Turvey, M. T. On the coordination and control of naturally developing systems. In J. A. S. Kelso & J. E. Clark (Eds.), *The development of movement control and coordination*. Chichester: John Wiley, 1982.

Lame, G. D. Is there phonology after universal phonetic interpretation? *Texas Linguistic Forum*, 1981, *17*, 45-62.

Liberman, A. M. On finding that speech is special. *American Psychologist*, 1982, *37*, 148-167.

Löfqvist, A., Baer, T., & Yoshioka, H. Scaling of glottal opening. *Phonetica*, 1981, *38*, 265-276.

MacKain, K. S., Best, C. T., & Strange, W. Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics*, 1981, *2*, 369-390.

Repp, B. H. Phonetic trading relations and context effects: New evidence for a phonetic mode of perception. *Psychological Bulletin*, in press.

Repp, B. H., & Mann, V. A. Fricative-stop coarticulation: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, in press.

Saltzman, E. L., & Kelso, J. A. S. Toward a dynamical account of motor memory and control. In R. Magill (Ed.), *Memory and control in motor behavior*. Amsterdam: North Holland, in press.

Shankweiler, D., Liberman, I. Y., & Mark, L. S. Phonetic coding in dyslexics and normal readers, by Hall, Ewing, Tinzmann, and Wilson: A reply. *Bulletin of the Psychonomic Society*, 1982, *19*, 78-79.

APPENDIX

DTIC (Defense Technical Information Center) and ERIC (Educational Resources Information Center) numbers:

Status Report		DTIC	ERIC
SR-21/22	January - June 1970	AD 719382	ED 044-679
SR-23	July - September 1970	AD 723586	ED 052-654
SR-24	October - December 1970	AD 727616	ED 052-653
SR-25/26	January - June 1971	AD 730013	ED 056-560
SR-27	July - September 1971	AD 749339	ED 071-533
SR-28	October - December 1971	AD 742140	ED 061-837
SR-29/30	January - June 1972	AD 750001	ED 071-484
SR-31/32	July - December 1972	AD 757954	ED 077-285
SR-33	January - March 1973	AD 762373	ED 081-263
SR-34	April - June 1973	AD 766178	ED 081-295
SR-35/36	July - December 1973	AD 774799	ED 094-444
SR-37/38	January - June 1974	AD 783548	ED 094-445
SR-39/40	July - December 1974	AD A007342	ED 102-633
SR-41	January - March 1975	AD A013325	ED 109-722
SR-42/43	April - September 1975	AD A018369	ED 117-770
SR-44	October - December 1975	AD A023059	ED 119-273
SR-45/46	January - June 1976	AD A026196	ED 123-678
SR-47	July - September 1976	AD A031789	ED 128-870
SR-48	October - December 1976	AD A036735	ED 135-028
SR-49	January - March 1977	AD A041460	ED 141-864
SR-50	April - June 1977	AD A044820	ED 144-138
SR-51/52	July - December 1977	AD A049215	ED 147-892
SR-53	January - March 1978	AD A055853	ED 155-760
SR-54	April - June 1978	AD A067070	ED 161-096
SR-55/56	July - December 1978	AD A065575	ED 166-757
SR-57	January - March 1979	AD A083179	ED 170-823
SR-58	April - June 1979	AD A077663	ED 178-967
SR-59/60	July - December 1979	AD A082034	ED 181-525
SR-61	January - March 1980	AD A085320	ED 185-636
SR-62	April - June 1980	AD A095062	ED 196-099
SR-63/64	July - December 1980	AD A095860	ED 197-416
SR-65	January - March 1981	AD A099958	ED 201-022
SR-66	April - June 1981	AD A105090	ED 206-038
SR-67/68	July - December 1981	**	
SR-69	January - March 1982	**	**

Information on ordering any of these issues may be found on the following page.

**DTIC and/or ERIC order numbers not yet assigned.

AD numbers may be ordered from:

U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Road
Springfield, Virginia 22151

ED numbers may be ordered from:

ERIC Document Reproduction Service
Computer Microfilm International
Corp. (CMIC)
P.O. Box 190
Arlington, Virginia 22210

Haskins Laboratories Status Report on Speech Research is abstracted in *Language and Language Behavior Abstracts*, P.O. Box 22206, San Diego, California 92122.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Haskins Laboratories 270 Crown Street New Haven, CT 06510	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP N/A

3. REPORT TITLE

Haskins Laboratories Status Report on Speech Research, SR-69, January-March, 1982

4. DESCRIPTIVE NOTES *(Type of report and, inclusive dates)*

Interim Scientific Report

5. AUTHOR(S) *(First name, middle initial, last name)*

Staff of Haskins Laboratories, Alvin M. Liberman, P.I.

6. REPORT DATE March, 1982	7a. TOTAL NO. OF PAGES 306	7b. NO. OF REFS 583

8a. CONTRACT OR GRANT NO. HD-01994 NS13870 HD-16591 NS13617 N01-HD-1-2420 NS18010 RR-05596 PRF-8006144 BNS-8111470	9a. ORIGINATOR'S REPORT NUMBER(S) SR-69 (1982) 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)* None

10. DISTRIBUTION STATEMENT

Distribution of this document is unlimited*

11. SUPPLEMENTARY NOTES N/A	12. SPONSORING MILITARY ACTIVITY See No. 8

13. ABSTRACT

This report (1 January-31 March) is one of a regular series on the status and progress of studies on the nature of speech, instrumentation for its investigation, and practical applications. Manuscripts cover the following topics:

--Speech perception and memory coding in relation to reading ability
--The use of orthographic structure by deaf adults: Recognition of fingerspelled letters
--Exploring the information support for speech
--The stream of speech
--Using the acoustic signal to make inferences about place and duration of tongue-palate contact
--Patterns of human interlimb coordination emerge from the properties of non-linear limit cycle oscillatory processes: Theory and data
--Motor control: Which themes do we orchestrate?
--Exploring the nature of motor control in Down's syndrome
--Periodicity and auditory memory: A pilot study
--Reading skill and language skill
--On the role of sign order and morphological structure in memory for American Sign Language sentences
--Perception of nasal consonants with special reference to Catalan
--Speech production characteristics of the hearing impaired

DD FORM 1 NOV 65 1473 (PAGE 1)

UNCLASSIFIED

Security Classification

S/N 0101-807-6811

*This document contains no information not freely available to the general public. It is distributed primarily for library use.

A-31408

UNCLASSIFIED
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Speech Perception:						
memory coding						
audio memory, periodicity						
place of articulation, contact duration						
nasal consonants, Catalan						
temporal coherence, sinewave synthesis						
Speech Articulation:						
point vowels, proprioception, jaw						
Reading:						
memory coding						
orthography, use by deaf, fingerspelling						
symbolic order, morphology, American Sign Language						
short-term memory, comprehension						
Motor Control:						
Down's syndrome						
interlimb coordination, limit cycle oscillation						
many-variable problem						

DD FORM 1 NOV 65 1473 (BACK)
S/N 0101-807-6821

UNCLASSIFIED
Security Classification

A-31409